

OPTIMIZED COGNITIVE TRAINING: INVESTIGATING THE LIMITS OF
BRAIN TRAINING ON GENERALIZED COGNITIVE FUNCTION

A Dissertation
Presented to
The Academic Faculty

by

Hillary Schwarb

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Psychology

Georgia Institute of Technology
May 2012

OPTIMIZED COGNITIVE TRAINING: INVESTIGATING THE LIMITS OF
BRAIN TRAINING ON GENERALIZED COGNITIVE FUNCTION

Approved by:

Dr. Eric H. Schumacher, Advisor
School of Psychology
Georgia Institute of Technology

Dr. Paul Verhaeghen
School of Psychology
Georgia Institute of Technology

Dr. Audrey Duarte
School of Psychology
Georgia Institute of Technology

Dr. Edward Awh
Department of Psychology
University of Oregon

Dr. Randall W. Engle
School of Psychology
Georgia Institute of Technology

Date Approved: March 7, 2012

This dissertation is dedicated to my father for his endless support, encouragement, and unwavering confidence in my potential. Also for nurturing in me a love for science and for teaching me that in this life there are no problems, only solutions waiting to be discovered. I am eternally grateful for his dedication and his love. I miss him terribly.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Eric Schumacher, for his constant support and encouragement over the years. Any successes I have had as a graduate student are direct reflections of his careful guidance. I am forever grateful for his willingness to allow me to investigate those questions that interest me the most and for always providing me with the necessary resources to complete my projects. While my skill as a researcher is undoubtedly the result of his influence, more importantly perhaps, he has instilled in me an excitement for the data, an enthusiasm for seeing the puzzle and putting together the pieces to discover something new, and for this I am eternally indebted to him. I would also like to thank the members of my committee, Ed Awh, Audrey Duarte, Randy Engle, and Paul Verhaeghan, for their thoughtful comments and suggestions that have transformed this project into a necessary contribution to the literature. I am so appreciative of their continued support and their confidence in my ability to succeed even from my first days at the Georgia Tech. Thank you to Tom Redick and Nate Parks for all of their expertise, guidance, and friendship. Thank you also to my lab mates, especially Erin Lightman, Keith Main, Brian Roberts, Jayde Nail, and Zain Sultan, for their willingness to always help in a pinch and for their pivotal role in data collection. Thank you to Edward Ester whose help and patience was pivotal in allowing me to collect and process data from the short-term recalls. Thank you to the dozens of devoted subjects who walked across campus in rainstorms and the heat of a Georgia summer to participate in this project. Without their commitment and reliability, this work would never have been possible. Thank you also to my family, my mother and

my brother, for their continued commitment to my education over the last 28 years, and for their love, support, and constant enthusiasm. Thank you to my roommates, Deborah Willingham and Givenski Rogers, and my neighbors, Shawn and Cara Fausset, for always being there to let the dogs out when I needed to stay late in the lab, feeding me when the other option was cereal (again), and for always making me laugh. And finally, this research was supported in part by a grant to the author from the American Psychological Association; in part by a grant from the Air Force Offices of Scientific Research; and in part by a graduate research award from the Georgia Institute of Technology School of Psychology.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	iv
LIST OF FIGURES	viii
LIST OF ABBREVIATIONS	ix
SUMMARY	xi
 <u>CHAPTER</u>	
1 INTRODUCTION	1
Working Memory Capacity	2
2 SKILL TRAINING	5
Generalized Lessons from Skill Training	5
Feedback	5
Distribution of Practice and Retention	6
Transfer	7
Recommendations	10
The Cognitive Neuroscience of Training	10
3 COGNITIVE TRAINING	14
Cognitive Training Transfer Successes	14
Cognitive Training Transfer Failures	19
4 VISUAL SHORT-TERM MEMORY	22
Visual Short-Term Memory Capacity	22
A New Conceptualization of Short-Term Memory Capacity	25
Training Visual Short-Term Memory Capacity	29
5 THE PRESENT EXPERIMENT	35

Method	38
Participants	38
Groups	38
Training Session Tasks	39
Battery Session Tasks	40
Analyses	46
Procedure	47
Results	48
Training Tasks	48
Battery Tasks	49
6 GENERAL DISCUSSION	64
The Training Task	64
Transfer of Cognitive Skill	67
Working Memory	67
Attentional Control	68
Fluid Intelligence	69
Visual Short-Term Memory	70
Contact Control Group	72
What, Then, Does Working Memory Training Train?	74
Explanation of Failed Transfer	75
Explanation of Successful Transfer	76
Cross-Trial Attentional Control	80
Conclusions	83
REFERENCES	84

LIST OF FIGURES

	Page
Figure 1: Two Stimulus Classes	27
Figure 2: (a) Training Task Data (b) Training Gain	49
Figure 3: Change Detection Task Data	50
Figure 4: (a) Change Detection Task Measures of Number (b) Resolution	51
Figure 5: (a) VT Group (b) ST Group (c) NCC Group VSTM Capacity Estimates	52
Figure 6: (a) Color Short-Term Memory Task Measures of Number (b) Resolution	53
Figure 7: (a) Spatial Short-Term Memory Task Measures of Number (b) Resolution	54
Figure 8: Visual Short-Term Memory Capacity Estimates across Tasks	55
Figure 9: (a) VSTM Composite Scores of Number (b) Resolution	57
Figure 10: (a) Operation Span Task Data (b) Symmetry Span Task Data	58
Figure 11: Working Memory Task Composite Scores	59
Figure 12: (a) RAPM Task Data. (c) CCF Task Data.	60
Figure 13: Fluid Intelligence Task Composite Scores	61
Figure 14: (a) Flanker Task Data (b) Antisaccade Task Data	62
Figure 15: (a) Optimized VT Comparison (b) Optimized ST Comparison	66
Figure 16: Flanker Task Sequential Effects	82

LIST OF ABBREVIATIONS

AC	Attentional Control
BS1	Battery Session 1
BS2	Battery Session 2
BOLD	Blood Oxygenation Level Dependent
BOMAT	Bochumer Matrizen-Test
CCF	Cattell Culture Fair
cf.	Compare
CDA	Contralateral Delay Activity
ERPs	Event Related Potentials
<i>Gf</i>	Fluid Intelligence
e.g.,	For Example
<i>F</i>	F-Statistic
fMRI	Functional Magnetic Resonance Imaging
Pmem	Measure Of VSTM Number
SD	Measure Of VSTM Resolution
vis.	Namely
NCC	No-Contact Control
η_p^2	Partial Eta Squared
<i>r</i>	Pearson Correlation Statistic
<i>p</i>	P-Value
RAPM	Raven's Advanced Progressive Matrices

ST	Spatial Training
i.e.,	That Is
<i>t</i>	T-Statistic
VT	Verbal Training
VSTM	Visual Short-Term Memory
WM	Working Memory

SUMMARY

Since antiquity, philosophers, theologians, and scientists have been interested in human memory; however, researchers today are still working to understand the capabilities, boundaries, and architecture. While the storage capabilities of long-term memory are seemingly unlimited (Bahrick, 1984), working memory, or the ability to maintain and manipulate information held in memory, seems to have stringent capacity limits (e.g., Cowan, 2001). Individual differences, however, do exist and these differences can often predict performance on a wide variety of tasks (cf. Engle, 2001). Recently, researchers have promoted the enticing possibility that simple behavioral training can expand the limits of working memory which indeed may also lead to improvements on other cognitive processes as well (cf. Morrison & Chein, 2011). The current study investigated this possibility. Recommendations from the skill training literature (cf. Schneider, 1985) were incorporated to create optimized verbal and spatial working memory training tasks. Significant performance improvements were evident across eight days of cognitive training using verbal and spatial adaptive n-back procedures. Training-related improvements were also evident for some untrained measures of visual short-term memory, attentional control, and working memory. These training effects, however, were not universal. Other measures of visual short-term memory and attentional control, as well as measures of fluid intelligence were unaffected by training.

CHAPTER 1

INTRODUCTION

Psychologists have been interested in understanding learning and training for over a century (Ebbinghaus, 1885/1913). Indeed William James (James, 1890), in his seminal *Principles of Psychology*, identified practice as an important area of research in understanding human behavior. Since the turn of the twentieth century, considerable research has been conducted in an attempt to understand how skill acquisition and performance improvements occur and to identify those conditions that promote successful learning and those that hinder it. Recently researchers have become interested not only in training behaviors and skills, but also in developing and strengthening the underlying cognitive processes that support behavior (see Morrison & Chein, 2011; Shipstead, Redick, & Engle, 2010 for a review). Popular culture and the scientific community alike have taken a particular interest in working memory (WM) training because WM appears to be a central component to critical real-world abilities such as fluid intelligence (e.g., Engle, Kane, & Tuholski, 1999; Kane, Conway, Miura, & Colflesh, 2007), mind wandering (Kane et al., 2007), and controlled attention (e.g., Conway, Cowan, & Bunting, 2001; Kane & Engle, 2003b; Unsworth, Schrock, & Engle, 2004) to name a few. Thus the potential benefit of enhancing WM capacity has seemingly limitless real-world applications.

WM is generally defined as the ability to actively maintain, monitor, and manipulate information in memory (Baddeley & Logie, 1999). It is generally believed that there are strict limits on the amount of information that can be held in WM at a given

time (e.g., Cowan, 2001; Miller, 1956). Cognitive training is one possible method in which we may be able to expand the limits of WM capacity. This may in turn enhance those abilities related to WM. The current study sought to investigate the influence of cognitive training on a variety of related cognitive processes, namely WM, attentional control (AC), fluid intelligence (*Gf*), and visual short-term memory (VSTM). The current study is novel in two ways. First, training recommendations from the skill training literature were applied to the cognitive training domain so as to create an optimal WM training design. Second, this study specifically addresses cognitive training efficacy on VSTM performance; an underinvestigated area in the brain training literature (c.f. Morrison & Chein, 2011).

Working Memory Capacity

The term WM, as previously noted, refers to the ability to actively maintain and manipulate information in memory (Baddeley & Logie, 1999). This ability is critically important to the successful performance of many, perhaps even most, of the tasks that individuals complete each day (Jonides et al., 2008). It is likely for this reason that understanding memory has been a primary pursuit of experimental psychologists since the inception of the field (e.g., James, 1890). Although the WM literature is vast and rich—indeed dozens of detailed models of working have been proposed (e.g., Baddeley & Logie, 1999; Cowan, 1999; Engle, et al., 1999)—for the present purposes, I focus here on one small aspect of WM, namely the limits of capacity.

In his seminal paper, Miller (1956) investigated the limits of memory by systematically studying unidimensional discrimination judgments across modalities (e.g., vision, audition, gestation, etc.). Investigating this wide variety of variables and tasks, he

determined that the range of items available in memory was remarkably similar: About seven plus-or-minus two. Given that memory capacity appeared reasonably constant across modalities, Miller began to identify techniques that would allow us to expand our memory capacity allowing us to hold more information in memory at one time. He determined that making relative instead of absolute judgments was one way to improve memory accuracy. Alternatively, he demonstrated that recoding the items into meaningful chunks of information can dramatically improve the limits of memory.

In a modern revisitation of this issue, Cowan (2001) conducted an extensive review of the literature and determined that capacity estimates may not be as optimistic as Miller (1956) had suggested. Cowan proposed four plus-or-minus one as the pure capacity limit of WM. He further suggested that overestimates in the literature were likely the result of a failure to control for rehearsal. He reported that when rehearsal was restricted by size of the memory item (e.g., famous phrases or proverbs), non-verbalizability of the memory item (e.g., unfamiliar Chinese characters), or articulatory suppression (e.g., repeating a single word continuously), capacity limits were consistently estimated at approximately four items.

Although most studies estimate that mean WM capacity is approximately four items (c.f., Cowen, 2001), considerable individual differences are reported in the literature and these differences are often predictive of performance on a wide variety of tasks (cf. Engle, 2001). For example, it has been demonstrated that individual differences in WM capacity predict performance on AC tasks such as the Stroop task (e.g., Kane & Engle, 2003a), the flanker task (e.g., Heitz & Engle, 2007) the antisaccade task (e.g., Unsworth, et al., 2004), and dichotic listening tasks (e.g., Conway, et al., 2001). It has

also been demonstrated that individual differences in WM capacity predict participants' ability to select task-relevant stimuli (Vogel, McCollough, & Machizawa, 2005), avoid attentional capture from irrelevant stimuli (Fukuda & Vogel, 2009), and recover from failures to ignore irrelevant stimuli (Fukuda & Vogel, 2011). Additionally, it has been demonstrated that individuals with low WM capacity show a higher proclivity toward mind wandering (e.g., Kane, Brown, et al., 2007). WM capacity has also been demonstrated to predict performance on higher-level reasoning tasks such as tests of general *Gf* (e.g., Kane, Conway, et al., 2007) as well as Scholastic Aptitude Test (SAT) performance (e.g., Turner & Engle, 1989). This list is only a small example of the vast variety of tasks for which individual differences in WM capacity can predict performance. Thus, it is no surprise that researchers are interested in developing methods to improve WM capacity which may then lead to improvements in other cognitive skills.

CHAPTER 2

SKILL TRAINING

Cognitive training or brain training, the notion that practicing some cognitive skill will result in improved performance in a variety of other frequently used skills in day-to-day life, has gained substantial popularity and enthusiasm both in popular culture and research domains. There are several companies who have capitalized on the possibility of general improvement from focused practice and have developed their own brain training programs (e.g., Cogmed, 2006). Such efforts, however, may be premature as scientific results on the efficacy of brain training are currently mixed (see Shipstead, Redick, & Engle, 2010 for a review). Before delving into the cognitive training literature in more detail, it is important to note that there is a substantial and rich skill training literature that has frequently been neglected when developing cognitive training experiments. This literature is undoubtedly important because there are many seemingly logical assumptions made about training (e.g., practice makes perfect, training should primarily focus on accurate performance, etc.) that are often entirely incorrect (Schneider, 1985). It is for this reason, that a careful review of the skill training literature, particularly concerning practice schedules and feedback, is important and can inform us how to best design cognitive training studies in the future.

Generalized Lessons from Skill Training

Feedback

Bartlett (1947) stated that simply practicing a skill will not lead to perfection of that skill, but rather an individual will only master a skill if feedback is given so as to

guide performance. An individual who is unaware of the results of his or her practice will not show marked improvements. Although research with rats has suggested that the immediacy of feedback is important for skill learning (Adams, 1987), research with humans has suggested that immediacy of feedback is not essential (Pashler, Rohrer, Cepeda, & Carpenter, 2007). The quality of the feedback, however, does seem to play an important role. Impoverished feedback (i.e., “right” or “wrong”) does little to improve performance, however, quantitative feedback (Adams, 1987) or providing the learner with the correct answer (Pashler, et al., 2007) can dramatically improve performance.

Distribution of Practice and Retention

Again it is worthwhile to note that several findings from training research appear counterintuitive. This is particularly apparent when considering distribution of practice and retention. A common finding in the literature is that poor performance during training may result in good retention performance whereas good performance during training results in poor retention performance (Schmidt & Bjork, 1992). As early as 1932, Edward Tolman asserted the importance of disentangling learning from performance (in Adams, 1987). One prevalent finding in the literature is that random practice produces poor training phase performance, but good retention compared to block practice (Rogers, 1996; Schmidt & Bjork, 1992). Shea and Morgan (1979) conducted a study requiring complex and sequenced motor movements. There were three types of trials that were either blocked (several trials of Task1 completed before moving on to Task 2, etc.) or random (all three trial types were intermixed in a given block of trials). Participants were assigned to either the Blocked or Random Group and instructed to respond as quickly as possible. During the acquisition phase, the Blocked Group consistently performed more

quickly than the Random Group. However, a retention test 10 days later demonstrated that regardless of whether participants were tested under blocked or random conditions, those participants who were in the Random Group during training performed better than those participants in the Blocked Group. This finding has been frequently replicated in the motor learning literature (e.g., Rogers, 1996; Schmidt & Bjork, 1992) as well as the verbal skill learning literature (e.g., Dempster, 1988; Pashler, et al., 2007; Schmidt & Bjork, 1992). These studies suggest that random practice schedules improve retention because they require additional information processing.

Another method for encouraging additional processing is inducing variability into a training program (Schmidt & Bjork, 1992). Kerr and Booth (1978) conducted a prototypical study demonstrating the importance of variability. They trained two groups of 8-year-olds to toss a beanbag through a hole. One group always practiced throwing the beanbag from a distance of three feet. The other group fluctuated between practicing throwing the beanbag from two or four feet. Accuracy of all participants was then tested from a distance of three feet. Perhaps surprisingly, the group who practiced under variable conditions consistently outperformed the group that consistently practiced at the testing distance. This result was also replicated with a group of 12-year-olds (Kerr & Booth, 1978). This experiment demonstrates that when additional processing is encouraged and deeper understanding is developed via a variable training procedure, performance improves.

Transfer

In this discussion the importance of the distribution of practice, it is apparent that the goal of training should not simply be good training phase performance, but rather

good learning and retention. Retention is certainly a pivotal component of training, however, perhaps more important is the issue of transfer or generalizability. In a real world setting, for example, it is unlikely that future instances of a trained skill will occur under identical conditions as the training procedure. A primary goal of training (i.e., educational, military or work-place training) is that learning will transfer outside of the classroom or training environment so that it can be applied in real-world situations (Adams, 1987; Barnett & Ceci, 2002; Schmidt & Bjork, 1992). This is, in fact, the premise upon which “cognitive training” is founded; increasing efficiency of general cognitive processes to improve performance in various aspects of life. Despite the importance of this goal, findings are mixed and many scientists have concluded that transfer rarely occurs and often fails (Thorndike & Woodworth, 1901; Detterman, 1993 in Barnett & Ceci, 2002). Others however have asserted that transfer can occur (Barnett & Ceci, 2002; Pashler, et al., 2007; Rogers, 1996; Schmidt & Bjork, 1992) and Barnett and Ceci have developed a taxonomy outlining the important components for successful far transfer.

Barnett and Ceci (2002) identified three content components (i.e., what) and six context components (i.e., when and where) that play an important role in skill transfer.

When designing a training program, there are three contextual factors to consider:

Learned skill (i.e., Is the skill a specific concept or procedure or is it more general such as a problem solving heuristic?); performance change (i.e., What performance component is intended to be transferred: speed, accuracy, or both?); and memory demands (i.e., Can the learned skill be prompted later or must an individual recognize when using the skill is appropriate?). General or deep principles are found to withstand far-transfer better than

specific concepts or procedures, so when possible, general rules or heuristics should be trained. Alternatively, successful transfer with and without memory demands seems to interact with the type of skill learned. Specific concepts transfer better when individuals are prompted, however, general rules show better transfer under conditions of spontaneous recall. Of the six context components, Barnett and Ceci identify knowledge domain, physical context, and temporal context as the most important. Knowledge domain is defined as the knowledge base to which the skill is most applicable; an example of far-transfer is applying a skill learned in a previous art class to a flight simulator task. Physical context refers simply to the physical location during learning and training; an example of far-transfer is learning a tactical maneuver in the classroom and then using that knowledge on the battlefield. Finally, temporal context simply refers to the amount of time that passes between training and testing phases; the hope is that transfer of any given skill will last for several years post-training.

It is difficult to make recommendations as to how to best control for these six components and thus ensure successful far-transfers because the studies currently available in the literature manipulate only one or a few components while holding the other components constant across conditions (Barnett & Ceci, 2002). Additionally, no studies have been reported to date demonstrating successful far-transfer across all or even most dimensions. This is no easy pursuit and future research is necessary to investigate the limits of transfer. Still, an understanding and awareness of the multidimensionality of transfer can help guide researchers to develop efficient training programs and make educated predictions about the potential for skill transfer.

Recommendations

In light of this vast body of literature that attempts to identify the various components that lead to successful training, Schneider (1985) outlined several rules for developing an advantageous training procedure that are still relevant today. He advised trainers to develop training programs that promote consistent processing and provide many trials of critical skills thus affording the learner many opportunities to process the task or a task component consistently. He suggested varying aspects of the task that would vary in the post-training environment and ensuring that learners remain motivated. He advised the training of strategies that minimize temporary memory components as well as practicing these strategies under mild speed stress and multi-task situations. Schneider asserted that careful consideration of these aspects of the training program would result in better learning of skills that would be readily applicable after training is completed. The main point thus far is that individuals can learn and can be trained and that those skills acquired during training can transfer to other situations. The caveat, however, is that training and transfer often fail and trainers need to take special care to ensure that the training environment is organized to promote transfer of the acquired skill. The current training design seeks to incorporate these recommendations from the skill training literature so as to optimize the efficiency of the training design.

The Cognitive Neuroscience of Training

With appropriate training programs, skilled performance can improve and just as behavioral changes can occur, so can the underlying neural circuits (e.g., Erickson et al., 2006; Posner, DiGirolamo, & Fernandez-Duque, 1997). This phenomenon generally termed “cortical plasticity.” More specifically, cortical plasticity refers to changes in

neural circuits (i.e., synaptic changes, the addition of new neurons, etc.) as a result of neural activity (Mercado, 2008). For many years it was believed that this sort of plasticity was reserved for children, however, recent studies have demonstrated that cortical plasticity can occur throughout one's lifetime (Mercado, 2008). In truth, researchers have only studied the neural changes that accompany practice for approximately ten years and therefore, there are a number of competing (while at the same time not mutually exclusive) theories for what happens to the brain with training.

Behavioral changes of a trained skill may or may not lead to detectable changes in brain activity. If brain activity does change with training, there are three dominant models concerning the nature of this change. The first model suggests that with training, increased activation may result as active brain regions increase in either size, number, or both (Petersen, van Mier, Fiez, & Raichle, 1998). A prototypical example of this expansion comes from Maguire and colleagues (2000) investigation of the brains of experienced, licensed London taxi drivers. They reported that these taxi drivers, who were highly skilled at navigating London roads without a map, had significantly larger posterior hippocampi than non-taxi-driving controls. Additionally, they demonstrated that hippocampus volume correlated with the number of years an individual had driven a taxi.

In addition to structural changes, other research has demonstrated that functional cortical expansion can occur with training. For example, Karni and colleagues (1995) asked participants to tap five-position sequences using the four fingers of their non-dominant hand as quickly as possible for 4-6 weeks. Participants were scanned once a week while performing both the trained sequence and a control sequence. Performance was evaluated for both the trained and control sequences across the experiment.

Behavioral data revealed that participants were significantly faster on the trained sequence at week five compared to week zero; however, they showed no significant improvement on the control sequence. Imaging data showed comparable activation in area M1 for both sequences at week 0; however, by week 5, the area of M1 activation was significantly larger for the trained than the control sequence. This study thus demonstrated functional cortical plasticity in the adult brain after a relatively short amount of time.

The second model states that with practice, new brain areas are engaged to perform task-related mental computations that were previously performed by other areas. This model predicts a circuit change as performance improves (Posner, et al., 1997). For example, Burton and colleagues (2002) used functional magnetic resonance imaging (fMRI) to identify active neural regions involved in Braille reading among the blind. They scanned individuals who were either congenitally blind or had late-onset blindness (after 12 years of age) during a verb generation task for Braille nouns. Regardless of the onset of a participant's blindness, all participants had been reading Braille for at least 12 years. All participants demonstrated significant visual cortex activation despite a total lack of vision. This activation included primary and secondary visual areas as well as some higher-level visual areas. Interestingly, however, only those individuals with congenital blindness showed significant activation in area V5/MT. This study again provides evidence that cortical reorganization can occur and can continue to develop even in adulthood.

The third model, called the scaffolding-storage framework, states that a set of scaffolding regions are active early in practice that support unskilled and effortful

performance. As proficiency develops, these novelty-related processes and consequently the regions supporting these processes are no longer necessary and drop out completely (e.g., Garavan, Kelley, Rosen, Rao, & Stein, 2000; Petersen, et al., 1998; Schumacher, Hendricks, & D'Esposito, 2005). With practice, processing becomes more efficient and in time task-related processes are accessed as rote programs supported by a separate set of regions. Peterson and colleagues investigated this idea using a verb generation task. In their experiment, participants were presented with a noun and were asked to produce a related verb. The presented nouns were either unpracticed (first presentation of a noun), practiced (same nouns repeated), or novel (new nouns presented after practice on the task); a control condition in which participants simply read the noun out loud was also included. Behavioral data indicated that with practice, verb generation becomes faster and the same verbs are used more frequently. PET results demonstrated that in both the unpracticed and novel conditions, there was an increase in neural activity in the left prefrontal cortex, anterior cingulate cortex and right cerebellum. Conversely, in the practiced and control conditions, only insula activation increased. These results suggest that a specific set of regions are active when a task is new and that these regions drop out over time and new more specific task-related regions come online. Peterson and colleagues reported similar results using a spatial maze-tracing task.

CHAPTER 3

COGNITIVE TRAINING

Given the apparent success of training on skill acquisition and evidence for neuronal effects of training, it is unsurprising that there has been an increased interest in the possibility that one can train various cognitive processes to improve cognitive performance more generally. The goal of cognitive training is twofold. First, brain training should improve the cognitive ability targeted by the training task and this improvement should endure over time (Willis et al., 2006). Second, these improvements acquired during training should transfer to other tasks and general everyday functioning (Willis, et al., 2006). As has previously been noted, transfer of trained skills to other tasks is infrequently reported in the literature (See Morrison & Chein, 2011 for a review). Transfer, however, can occur in certain situations.

Cognitive Training Transfer Successes

Generally, cognitive training experiments are designed so that participants complete a battery of tasks on the first day (i.e., battery session 1; BS1). Some or all participants return for some form of cognitive training that lasts several days, weeks, or months. Finally, all participants return to the lab and repeat the tasks completed during BS1 (i.e., battery session 2; BS2). In recent years, the number of studies adhering to this basic design has risen dramatically. However, despite the basic design similarities, these studies vary dramatically in other important features such as the duration of training, the training task used, the battery tasks investigated, and the comparison groups selected.

Outlined below are several studies that have demonstrated successful transfer across a variety of cognitive skills.

Green and Bavelier (2003, 2006) investigated the possibility that training on action-video-games can increase attentional capacity and that this increased attentional capacity transfers to a variety of other visual skills. Comparing gamers to nongamers, Green and Bavelier (2003) demonstrated that gamers show constant compatibility effects in the flanker task under all levels of difficulty. Nongamers, however, only showed a compatibility effect when the task was easy; when the task was difficult, the compatibility effect vanished. This suggests that gamers have an increased attentional capacity allowing them to process more information under greater demands. Gamers additionally demonstrated advantages in an enumeration task as well as attentional blink and useful-field-of view tasks. Comparing these naturally occurring groups, however, opens up the possibility of alternative explanations. Perhaps gamers enjoy video games and play them frequently because they simply have a higher propensity for visual task performance. In this situation, performance improvements on other visual tasks may not be attributable to video game training, but instead some innate ability common among gamers. To rule out this alternative explanation, Green and Bavelier trained one group of nongamers on an action-video-game (Medal of Honor; very attentionally demanding requiring simultaneous processing of multiple items) and another group of nongamers on Tetris (requires participants to only focus on one item at a time) for an hour each day for two weeks. All participants completed an enumeration, useful-field-of-view and attentional blink task before and after training. The participants who trained on the action-video-game consistently performed significantly higher after training than those

participants who trained on Tetris. Thus it seems that the attentional skills acquired during action-video-game playing generalize to detectable effects on new tasks, as well as performance in untrained locations. In a similar study, however, Boot and colleagues (2008) failed to find training-related improvement on measures of attention, memory, and executive control in a group of nongamers after 20 hours of action-video-game playing.

Jaeggi and colleagues (2008) explored the possibility that *Gf* could be enhanced by training participants on other cognitive tasks. According to Jaeggi and colleagues, *Gf* is “the ability to reason and solve new problems independent of previously acquired knowledge” (p. 6809). In the Jaeggi experiment, participants were trained on a dual *n*-back task in which participants were simultaneously presented with two strings of stimuli (letters and spatial locations) and asked to determine whether the current stimuli matched the stimuli that occurred *n* trials previously. The dual *n*-back task was adaptive (*n* changed based on performance) and thus a very demanding WM task. Training lasted between 8 and 19 sessions. Results revealed that participants who were trained with the dual *n*-back task showed significant improvement on measures of *Gf* (i.e., measured either with the Raven’s advanced progressive matrices task [RAPM] or the Bochumer Matrizen-Test [BOMAT]) from pretest to posttest. A control group that did not receive training (i.e., no-contact control (NCC) group) showed no such improvement. These data were later replicated using a single *n*-back training task (Jaeggi et al., 2010). Thus it seems that cognitive skill acquisition resulting from training on a demanding WM task transfers to situations requiring enhanced *Gf*. It should be noted, however, that these data have been recently criticized for non-trivial design flaws and inappropriate analyses (Moody, 2009; Morrison & Chein, 2011; Redick et al., 2011; Shipstead, et al., 2010).

Chein and Morrison (2010) investigated the efficacy of WM training on both near and far transfer by training participants on both spatial and verbal complex WM span tasks. After 20 training sessions, participants showed improvement on a composite memory measure that included both the trained tasks and spatial and verbal short-term memory tasks (compared to a NCC group). Participants also showed training-related improvements on the Stroop task (i.e., measure of AC) and the Nelson Denny reading comprehension task. These results provide evidence for successful near transfer of WM training to measures of WM, short-term memory, and AC as well as far transfer to measures of reading comprehension.

Finally, data from our laboratory (Schwarb et al., under review) evaluated the effects of WM training on a variety of untrained tasks. We trained participants on both spatial and verbal versions of an adaptive *n*-back task for 8 one-hour training sessions. Before and after training participants completed the automated operation span task (i.e., measure of verbal WM), the automated symmetry span task (i.e., measure of spatial WM), RAPM (i.e., measure of *Gf*), and the change detection task (i.e., measure of VSTM number and resolution). Compared to a NCC group, participants showed significant training-related improvements on the automated operation span task and most interesting for the purposes of the present study, we demonstrated significant training-related improvements on measures of VSTM resolution the change detection task and a trend toward a significant training-related improvement on measures of VSTM number. Unlike the Jaeggi studies, however, participants failed to show a training-related improvement on RAPM. These data suggest that while adaptive *n*-back training does not result in global improvement, training can improve some measures of WM as well as VSTM.

However, this study did not include a contact control group, so alternate interpretations of these data are possible (Shipstead, et al., 2010).

All of the studies reported thus far indicate the potential of generalized training programs to improve a variety of cognitive abilities. There are other researchers, however, that believe that transfer is likely more localized and that when a specific cognitive process is trained, transfer to tasks requiring that cognitive process is possible, but transfer to other different cognitive tasks is less likely.

Dahlin and colleagues (Dahlin, Neely, Larsson, Backman, & Nyberg, 2008) provide a good example of the limits of cognitive training and transfer. In their study, participants were trained on an updating task (i.e., remembering strings of letters) and then asked to perform a 3-back task (which requires updating) and a Stroop task (no updating required). After five weeks of training on the updating task, participants performed significantly better on the 3-back task compared to participants who were not trained. Post-training performance on the Stroop task, however, did not differ between the two groups. Thus training on a component skill required for the 3-back task (i.e., updating) significantly improved performance on the 3-back task. Dahlin and colleagues (Dahlin, Nyberg, Backman, & Neely, 2008) later replicated the finding that training on updating tasks improves performance on an untrained 3-back task and further demonstrated that the benefit of training on the updating tasks persisted 18 months after training ended.

Similarly, Li and colleagues (2008) trained participants on a standard spatial 2-back task and a spatial 2-back task that required participants to rotate all stimuli one position to the right. Participants completed 45 15-minute training sessions. Participants

showed a significant training-related improvement on an untrained spatial 3-back task, an untrained verbal 2-back task, and an untrained verbal 3-back task compared to participants in a NCC group. No group differences were evident, however, on either the rotation span or operation span tasks (i.e., complex span tasks measuring WM capacity) indicating that training-related improvements were only evident in near-transfer situations.

Taken together these studies suggest that cognitive training can indeed improve cognitive performance. However, such cognitive training is most effective if the important component processes of the transfer task are also included in the training tasks (Dahlin, Neely, et al., 2008; Dahlin, Nyberg, et al., 2008; Li, et al., 2008). Thus, the careful selection of training tasks is likely an important contributor to successful transfer.

Cognitive Training Transfer Failures

The findings reported above seem to indicate that cognitive training is effective in numerous situations. However, there is a growing literature that suggests that the breadth of training-related improvements on untrained cognitive tasks is selective at best (cf. Morrison & Chein, 2011; Shipstead, et al., 2010) and that despite the enthusiasm surrounding cognitive training, evidence for the efficacy of such programs is inconsistent (Owen et al., 2010). While several studies have provided evidence for training-related improvements on a variety of untrained tasks, these same studies, as well as additional studies have reported no such improvements on other similar and sometimes identical untrained tasks (c.f. Morrison & Chein, 2011). Outlined below are several findings that demonstrate a lack of training-related transfer to a variety of cognitive skills.

As noted previously, while Dahlin and colleagues (Dahlin, Neely, et al., 2008; Dahlin, Nyberg, et al., 2008) and Li and colleagues (Li, et al., 2008) demonstrated successful training-related improvements on untrained cognitive tasks, this only occurred for near-transfer tasks and there was no evidence for training-related improvements on far-transfer tasks (i.e., Stroop, WM span, letter and category fluency, and explicit recall tasks). Similarly, Jaeggi and colleagues (2008) demonstrated that dual *n*-back training improved measures of *Gf*, but failed to transfer to improvements on the Stroop task. Additionally we recently demonstrated that while spatial and verbal adaptive *n*-back training shows positive transfer to measures of verbal WM and VSTM (Schwarb, et al., under review), we failed to demonstrate any training-related improvement on other measures spatial WM (i.e., automated symmetry span task), *Gf* (i.e., RAPM task), or measures of focused attention (i.e., motion interference task and rapid decision making task).

Thus it is evident that results both across studies and within studies in the cognitive training literature are mixed concerning the efficacy of cognitive training on untrained tasks. Recently, Shipstead, Redick, and Engle (2010) have argued that one possible limitation in the literature to date is the inappropriate use of NCC participants as a comparison group. When only a NCC group is used to evaluate the efficacy of training, there is the potential for numerous alternative explanations of the resulting data. For example, training-related improvements could be the product of placebo, Hawthorn, or simple motivational effects rather than the strengthening of component cognitive processes through training.

In fact, two recent cognitive training studies using a contact control group to evaluate training-related improvements report a total absence of training-related improvement on untrained tasks. For example, Owen and colleagues (2010) trained 11,340 participants over an average of 24 training sessions. Participants were assigned to three groups: One group performed reasoning, planning, and problem solving tasks during training (experimental group 1), a second group performed memory, attention, visuospatial processing, and mathematical calculation tasks during training (experimental group 2), and a final group performed an obscure knowledge trivia-type task during training (contact control group). There were no significant differences on measures of grammatical reasoning, verbal short-term memory, spatial WM, or paired associate learning among the groups indicating no effect of training. Similarly, Redick and colleagues (2011) trained participants on either an adaptive dual n -back task or an adaptive visual search task over the course of 20 training sessions. Data indicate no difference between the two training groups and a NCC group on multiple measures of Gf , WM capacity, crystallized intelligence, perceptual speed, or multitasking abilities.

It is therefore evident that the literature is inconsistent regarding the efficacy of cognitive training on transfer to novel tasks. Many studies report evidence of training-related improvement on untrained cognitive tasks, however, these data too are often inconsistent and results vary across experiments. Other studies report a total lack of evidence that cognitive training is effective in improving performance on untrained tasks. These conclusions demonstrate the importance of careful experimental design so as to maximize the possibility of training-related improvements across tasks and to minimize the potential for alternative explanations.

CHAPTER 4

VISUAL SHORT-TERM MEMORY

This review thus far has demonstrated that while a vast multitude of cognitive processes have been assessed in the cognitive training literature (e.g., WM, *Gf*, AC, multitasking ability, reasoning, and episodic memory to name a few), one area that has been underinvestigated to date is the efficacy of cognitive training on VSTM processes. Therefore, one of the goals of the current study is to address the issue of cognitive training on VSTM improvement. However, before VSTM training is considered, it is first important to discuss VSTM more generally and explore its architecture.

Visual Short-Term Memory Capacity

Consistent with Cowan's (2001) observations, typical estimates of VSTM are restricted to three or four items (e.g., Luck & Vogel, 1997; Pashler, 1988; Sperling, 1960). Although overall limits are generally agreed upon, there is considerable disagreement about whether these capacity limits are defined by the number of items held in memory (e.g., Awh, Barton, & Vogel, 2007; Luck & Vogel, 1997) or the complexity of those items (e.g., Alvarez & Cavanagh, 2004; Eng, Chen, & Jiang, 2005).

In an early attempt to characterize VSTM, Luck and Vogel (1997) conducted a series of change detection experiments demonstrating capacity estimates of approximately four items regardless of object complexity. In one experiment, participants were presented with colored horizontal bars at varying orientations. Participants were asked to remember either the color, orientation, or both color and orientation of the items briefly presented in a memory array. After a short delay, the memory array reappeared

and one of the items was cued. Participants determined whether the cued item was identical to the original memory display or whether that item had changed. Participants were equally adept at identifying changes to the display regardless of whether they were focusing on a single feature (color or orientation) or a conjunction of features (color and orientation). Vogel and Luck concluded that single-feature items were no better remembered than more complex multi-feature items. They extended this finding by noting that capacity estimates remained constant with four-feature items (e.g., color, orientation, size, and gap) compared to two- or single-feature items.

VSTM capacity estimates of approximately four items are also supported in the neuroscience literature. Indeed, data suggest a possible neural basis for such capacity limits. Vogel and Machizawa (2004) conducted a series of change detection experiments in which they cued participants to one hemifield and presented them with a bilateral memory array (colored squares). After a brief retention interval, a test array appeared and participants were asked to determine whether or not a change had occurred. Event related potentials (ERPs) were recorded and 200ms after the memory array, a large negative-going voltage was elicited in the contralateral hemisphere to the cued hemifield (termed contralateral delay activity or CDA). Interestingly, this CDA amplitude (focused over the posterior parietal and lateral occipital electrode sites) was found to be highly sensitive to the number of items in the memory array and reached asymptote at around four items. Thus CDA amplitude appears to track VSTM capacity. In a similar fMRI change detection experiment, Todd and Marois (2004; Experiment 1) presented participants with memory arrays composed of colored circles. After a brief retention interval, participants were presented with a single colored circle and asked whether or not this circle matched

the color of the circle presented in the same spatial location in the memory array. Todd and Marois reported a systematic increase of the bold oxygenation level dependent (BOLD) signal in the bilateral intraparietal and intraoccipital sulci until a set size of 3 or 4 after which the BOLD signal leveled off. Again, these data provide evidence that VSTM capacity limits are reflected in the brain.

It is potentially important to note, however, that both Vogel and Machizawa (2004) and Todd and Marois (2004) used very simple stimuli in their memory arrays and although both groups demonstrated neural correlates for capacity limits of about four items, other researchers suggest that the complexity of the stimuli may restrict capacity estimates (Alvarez & Cavanagh, 2004; Eng, et al., 2005). For example, Alvarez and Cavanagh (2004) performed a change detection experiment similar to Vogel and Machizawa (2004) except that they used five different classes of stimuli ranging in complexity: Colored squares, letters, Chinese characters, random polygons, and shaded cubes. They found reductions in VSTM capacity estimates as stimulus complexity (as determined by visual search slope) increased. Data showed capacity estimates of about 3.6 items for colored squares but only 1.7 items for shaded cubes. Similarly, in a series of ERP experiments, Luria and colleagues (2009) demonstrated that CDA amplitude was larger for complex versus simple objects indicating that VSTM capacity depends on stimulus complexity such that more capacity is required to maintain complex objects compared to simple objects. Indeed Gao and colleagues (2009) have replicated the Luria findings using simple and complex polygons as well as colored landolt rings. However, there are also discrepant findings regarding this result. For example, Perez and colleagues (cited as in preparation in Fukuda, Awh, & Vogel, 2010) conducted a similar experiment

using colored squares and abstract shapes. The resultant set size dependent CDA amplitudes were indistinguishable for these two sets of stimuli.

Nevertheless, neuroimaging data also exists to support a distinction between VSTM capacity and item complexity. Xu and Chun (2006) varied the complexity of stimuli in an fMRI change detection experiment. They identified dissociable regions of activation for the number of items stored in memory and the complexity of those items. As in the Todd and Marois (2004) study, Xu and Chun reported that inferior intraparietal sulcus activity systematically increased from set size one to four and then leveled off regardless of stimulus complexity. Activity in superior intraparietal sulcus and lateral occipital complex, however, followed this pattern only for simple stimuli. Complex stimuli resulted in elevated activity that was not modulated by set size. They concluded that both fixed number of items and the complexity of those items contribute to estimates of VSTM capacity. Furthermore, Fougny and colleagues (2010) demonstrated that while the capacity estimate for simple versus complex objects is similar, the precision with which these items are remembered is impaired as stimulus complexity increases. Thus there is controversial evidence regarding the importance of complexity to VSTM capacity estimates in the imaging literature.

A New Conceptualization of Visual Short Term Memory Capacity

Awh and colleagues (2007) have developed a two-factor hypothesis of VSTM that may provide a unifying account for the data described above. They argue that change detection performance depends both on the number of items held in WM, and the resolution or discriminability with which those items are stored. According to this account, VSTM capacity is similar for both simple and complex stimuli. Lower capacity

estimates for complex stimuli result from comparison errors between highly similar complex items (e.g., Awh, et al., 2007; Fukuda, Vogel, Mayr, & Awh, 2010; Scolar, Vogel, & Awh, 2008). Awh, Barton, and Vogel (Experiment 2) tested this hypothesis by modifying the type of change that occurred on each trial in the change detection task. Participants were presented with two classes of stimuli (shaded cubes and Chinese characters) on each trial. On each “change” trial, the change occurred either within-class (i.e., one shaded cube to another) or between-class (i.e., shaded cube to Chinese characters). Color stimuli were presented on other trials as a baseline measure for capacity estimates with simple stimuli. For within-class change trials, capacity estimates were low (1.4 for cubes and 1.7 for Chinese characters; replicating Alvarez & Cavanaugh, 2004) and equivalent to estimates from more traditional change detection studies (e.g., Awh, Barton, & Vogel, 2007; Experiment 1). However, for between-class change trials, capacity estimates were as high for complex items (4.2 for shaded cubes and 3.5 for Chinese characters) and simple items (3.6 for colors). These data suggest that the number of items stored in VSTM is similar regardless of stimulus complexity and that reduced capacity estimates are likely the result of comparison errors due to similarity among complex items. Furthermore, capacity estimates for highly distinguishable stimuli (i.e., between-class trials and colors trials) were highly correlated whereas capacity estimates for difficult to distinguish stimuli (i.e., within-class trials) and colors showed no correlation. Thus Awh and colleagues concluded that the number of items in WM and the resolution with which these items are stored are independent and dissociable processes that both contribute to in VSTM.

Further support for the two-factor hypothesis comes from Fuduka, Vogel, Mayr, and Awh (2010) who asked participants to perform a change detection experiment with two distinct classes of stimuli (rectangles and ovals) each with two different specific stimuli (see Figure 1; these same stimuli were used in the current study). A color stimulus comparison condition was included as in Awh, Barton, and Vogel, 2007. An exploratory factor analysis revealed that within-class changes and between-class changes loaded onto two orthogonal factors with no significant cross loadings. These findings support the two-factor hypothesis of VSTM capacity. Furthermore, confirmatory factor analysis comparing within-class changes, between-class changes, and Gf (measured by the RAPM and the Cattell Culture Fair [CCF] tests) revealed that the between-class factor and Gf were correlated ($r = .66$) while the within-class factor was not correlated with Gf ($r = -.05$). These data suggest that the relationship between WM capacity and Gf is mediated by the maximum number of items held in memory alone further supporting two separable processes involved in VSTM.

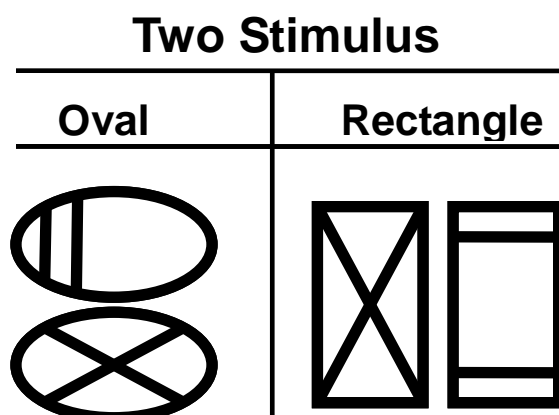


Figure 1. Two Stimulus Classes

The change detection task is not the only task that has been used to demonstrate separable number and resolution measures. Short-term recall tasks are also an effective means of separating these processes (Zhang & Luck, 2008). For demonstrative purposes, consider the color version of the short-term recall task presented in the original experiment (Zhang & Luck, 2008). In the color short-term recall task, on a given trial participants are briefly (e.g., 100ms) presented with some number of to-be-remembered colored squares. After a short delay (e.g., 900ms), one of the squares is probed and participants must indicate via mouse click on a color wheel, what color that square had been. The probability that a probed item was held in memory can be calculated by comparing the nearness of the recalled color to the actual color. Using standard estimation methods, this information can be used to extract measures of both number (i.e., Pmem component) and resolution (SD component). This task has been used to support the idea that VSTM capacity is restricted by a limited set of slots that can hold a limited set of representations (e.g., Zhang & Luck, 2008, 2011).

Finally, as with other forms of WM, there is a wide range of individual differences in VSTM capacity; in one study, for example, capacity estimates ranged between 1.7 and 6.4 items (Todd & Marois, 2004). The cognitive neuroscience literature suggests that these behavioral limits are reflected in the brain. For example, Todd and Marois (2005) used fMRI to demonstrate that neural activity in the intraparietal and intraoccipital sulci at a participant's individual capacity limit was correlated with the number of items stored in memory. In fact, they determined that individual load-modulated activity in this region accounted for about 40% of the variance in overall VSTM capacity estimates. With such a wide range of individual differences, it seems

plausible that it may be possible to improve VSTM capacity for some if not all participants.

In fact, there is some evidence that expertise with the stimuli is one way in which participants' can expand their VSTM capacity. For example, Curby and Gauthier (2007) reported larger capacity estimates for upright faces (a category for which all participants were experts) compared to inverted faces. Scolari, Vogel, and Awh (2008) investigated whether comparison errors could explain these findings by comparing VSTM capacity for within-class and between-class changes using upright faces, inverted faces, and shaded cubes. Results indicate that for between-class changes (i.e., upright faces to shaded cubes and inverted faces to shaded); capacity estimates were similar for upright and inverted faces. These data suggest that there is no expertise advantage when fine grained comparisons are not necessary. Importantly, participants showed a large advantage for detecting within-class changes for upright compared to inverted faces. Scolari and colleagues concluded that while expertise does not influence VSTM capacity, resolution depends largely on expertise. Thus it appears that expertise with the stimuli affects the success of VSTM resolution, but not capacity.

Training Visual Short-Term Memory Capacity

As with the cognitive training literature in general, the training literature regarding VSTM capacity is mixed with some researchers reporting training-related capacity improvements and other researchers reporting no improvement. The VSTM capacity training literature is diverse and the tasks used to measure capacity (e.g., change detection task vs. spatial span task) and the amount of training (e.g., 130 trials vs. 20+ days) varies widely. One important distinction between the VSTM training literature and

the cognitive training literature more generally is that in almost all instances, at least one of the training tasks used is a VSTM task. Thus unlike the general cognitive training literature, any training-related improvement on measures of VSTM constitutes near-transfer. Additionally, by including the assessment task as part of the training battery, it is impossible to disambiguate improvements due to training and improvements associated with simple familiarity with the task.

Still, there are several studies that report training-related behavioral improvements in VSTM capacity. For example, Klingberg and colleagues (2002; Experiment 2) gave participants an identical pre- and post-training battery of tasks including a spatial span task (i.e., measure of VSTM), a span board task (i.e., measure of VSTM), the Stroop task, RAPM, and a choice RT task. Participants returned to the laboratory for approximately 26 training sessions over the course of five weeks where they practiced several tasks: The spatial span task used during the battery sessions (i.e., measure of VSTM), a backwards digit-span task, a letter-span task, and a go/no-go type task. All training tasks included 30 trials and were adaptive (i.e., the difficulty increased as performance improved). Participants showed significant improvement on all training tasks as well as the span board task and RAPM. Peculiarly, improvements for this group of normal adults were compared to test-retest scores obtained in a population of ADHD children (Experiment 1). All improvements were greater for the experimental group than for the comparison group and the authors conclude that training was effective. However, given the differences between groups, these data should be interpreted with caution.

Beck and colleagues (2008) conducted a series of experiments demonstrating situations in which training-related improvements in VSTM capacity occur and situations

in which they do not. Participants completed 120-130 change detection trials with different stimulus types (objects vs. spatial locations). Beck and colleagues were particularly interested in participants' abilities to learn about change probabilities; however, the data are still interesting from a VSTM capacity perspective. After less than 15 min of practice, participants showed performance improvements for changes in spatial location, but not object identity. Thus these studies provide evidence that VSTM can improve or not depending on the type of stimuli used.

In addition to Beck and colleagues (2008), several other researchers have reported no improvement of VSTM capacity after training (e.g., Eng, et al., 2005; Olson & Jiang, 2004; Olson, Jiang, & Moore, 2005). For example, Olson and colleagues investigated whether repeating memory sets would improve performance in the change detection task. When the location of a target varied with each repetition of the memory set, participants did not show evidence for improved change detection performance with practice (Olson, Jiang, & Moore, 2005; Experiment 1). In fact, performance on novel memory sets did not differ from performance on repeated memory sets leading Olson and colleagues to conclude that training-related improvements did not occur. It is important to note, however, that when trials were binned into epochs and accuracy was compared across the experiment, the main effect of epoch was significant. Therefore change detection performance did improve during the experiment, but improvement was not limited to repeated memory sets. Also to note, in this study participants only completed 384 trials in a single session which is considerably less than 20-25 sessions of practice completed by participants in the Klingberg and colleagues (2002) and Olesen and colleagues (2004) studies. Thus it is evident that the behavioral data investigating the possibility of VSTM

capacity improvements are mixed. Some studies report training improvements while others do not; however, there are many clear differences between the tasks used to evaluate VSTM capacity as well as the number of trials that constitute “training.” There is, therefore, a need for additional carefully designed training studies that address the issue more directly.

In addition to behavioral data, imaging studies are also informative about the possibility of training-related performance improvements in VSTM capacity. Klingberg, Forssberg, and Westerberg (2002) investigated the neural correlates of VSTM capacity improvements accompanying development. VSTM capacity was determined using a spatial simple-span task with set sizes between 3 and 9 items. Participants were then scanned while performing a change detection type task where a series of 3-5 spatial locations were identified one at a time on a 4 x 4 matrix (i.e., measure of VSTM). After a brief delay, a target location was cued and participants had to decide whether or not one of the remembered locations matched the cued location. VSTM capacity correlated with neural activity in the left superior frontal sulcus and left intraparietal cortex. While this study does not address training-related VSTM capacity improvements, it does demonstrate that there are regions of the frontal and parietal cortex that activate differentially as VSTM capacity improves (though in this particular study, capacity improvements were the result of age and not training).

Olesen, Westerberg, and Klingberg (2004; Experiment 2) conducted the first study to investigate the effect of visual WM training on VSTM (note that while some of the studies reported previously use the term “visual WM” to describe the tasks used, for the purposes of this dissertation “WM tasks” are identified as tasks that include both a

storage and a processing (i.e., maintenance and manipulation) component as outlined by Baddeley and Loggie (1999)). To date, this is the only study that has broached the issue of far-transfer of cognitive training to measures of VSTM. In this study, participants completed a pre- and post-training battery of tests including the span board task (i.e., measure of VSTM), the Stroop task, the digit span task, and a verbal WM test. Participants completed 23 training sessions where they practiced 30 trials of each the grid, grid rotation, and 3D grid tasks (i.e., measures of visual WM) from the commercially available Cogmed training software. On days 2, 3, 8, and 23 of training, participants also completed a spatial simple-span task (i.e., measure of VSTM) while in the fMRI scanner. Participants showed significant improvement on all training tasks across the experiment. Also, participants showed a significant training-related improvement on the Stroop task as well as a trend toward significant training-related improvements on both the span board and digit span tasks compared to test-retest scores from a naïve group of participants. This study provides some preliminary evidence that WM training improvements may transfer to improvement on measures of VSTM (e.g., the span board task). It is important to note, however, that the number of participants included in these studies was very small ($N = 3$ in experiment 1 and $N = 8$ in experiment 2) so further research is necessary.

While many of these studies demonstrate enhanced performance on various VSTM tasks after training, this literature is deficient in at least three ways. First, none of these studies assess the possibly differential role of number and resolution processes contributing to VSTM capacity more generally. Second, while claims are made about the efficacy of training for enhancing VSTM capacity, the data only show post-training

performance improvements on VSTM tasks and do not measure the capacity estimates (e.g., Cowan's K) directly. Finally, because in most of these studies, the assessment tasks are also included as training tasks, it is impossible to dissociate performance improvements resulting from increased capacity or performance improvements resulting from some other process or processes in memory (e.g., familiarity with the task). The current study addresses these issues directly by comparing pre- and post-training number, resolution, and general capacity estimates in multiple VSTM tasks.

CHAPTER 5

THE PRESENT EXPERIMENT

At this point, several literatures have been summarized: The skill training literature, the cognitive training literature, and the VSTM capacity literature. The current study sought to apply the recommendations from both the skill training and cognitive training literatures to investigate the extent of adaptive *n*-back training improvement transfer to other cognitive tasks with a particular focus on the various components of Awh et al.'s (2007) two-factor model of VSTM capacity. Behavioral data suggests that VSTM number and resolution are separable processes (e.g., Awh, et al., 2007; Fukuda, Vogel, et al., 2010; Scolari, et al., 2008) both of which contribute to performance on the change detection task. There is some indirect evidence that distinct neural areas may mediate these processes (e.g., Ester, Serences, & Awh, 2009; Serences, Ester, Vogel, & Awh, 2009; Xu & Chun, 2006). Additionally, it has been demonstrated that neither number nor resolution measures can be enhanced with motivation, either instructional or monetary (Zhang & Luck, 2011). Finally, our laboratory has recently demonstrated that cognitive training can improve VSTM resolution and may influence VSTM number as well. However, given limitations of the experimental design (e.g., lack of a contact control group) alternate interpretations are possible (Shipstead, et al., 2010) thus necessitating further investigation.

Thus the aims of the current experiment were twofold. First, this study sought to develop an optimized training design by compiling all of the recommendations from the skill training and cognitive training literatures. Optimizing the training design should

promote performance improvements both on the training tasks themselves, but also on other untrained tasks and should additionally minimize the number of alternative explanations for the resulting data. Second, while this study sought to investigate the efficacy of cognitive training on a variety of cognitive functions, a particular interest was given to the role of cognitive training in improving VSTM performance. To date, while several studies demonstrate that VSTM training enhances VSTM task performance, only one study has investigated whether training other cognitive processes can transfer to VSTM (Olesen, et al., 2004) and possibly due to the small number of participants, trends in the data were not significant.

A wide variety of training designs have been used to study training-related cognitive enhancement; and perhaps consequently a wide variety of data both in support of and in opposition to training-related improvements have been reported. In order to optimize the training design, the current experiment carefully incorporated recommendations from both the skill training and cognitive training literatures (Adams, 1987; Bartlett, 1947; Pashler, et al., 2007; Redick, et al., 2011; Schneider, 1985; Shipstead, et al., 2010). First, participants were given a substantial number of training trials (Schneider, 1985). Training participants completed 4,800 scored trials and an average of 5,920 trials requiring a response. In the literature, the number of training trials varies dramatically across studies ranging from 270-10,000 trials per experiment with a mean of approximately 2,900 trials. Next, task requirements were varied (Schmidt & Bjork, 1992; Schneider, 1985) by training adaptive versions of both the spatial and verbal *n*-back tasks. To ensure that training remained variable across each training session, task difficulty was increased when participants spent more than five blocks at the same level

of difficulty. Of the 22 training studies reported here, 14 included an adaptive design and none controlled for static task requirements resulting from continual moderate performance. Additionally, feedback was given (Bartlett, 1947) during the experiment, however, it was given only at the end of each block of trials (Pashler, et al., 2007). All feedback was quantitative including both mean RTs and accuracy (Adams, 1987). Ten of the 22 training studies reported providing feedback (i.e., accuracy or accuracy and average RT), 5 of which provided feedback at the end of each block. Finally, participants were motivated throughout the study (Schneider, 1985). To this end, the experimenter provided verbal feedback at least every five blocks throughout the experiment and a monetary bonus was given for improved performance. Three of the 22 cognitive training studies reported here included a monetary or prize bonus based on performance. Finally, in addition to a NCC group, two experimental groups were included in the present study each serving as a contact control group for the other (Morrison & Chein, 2011; Shipstead, et al., 2010). Eight of the 22 reported training studies included a contact control group. Lastly, multiple tasks were included to index underlying cognitive abilities so as to enhance the generalizability of these findings (Morrison & Chein, 2011; Shipstead, et al., 2010). These recommendations were applied so as to include the most efficient training design possible.

Finally, to effectively evaluate the impact of cognitive training on measures of VSTM and to identify differential training effects on both number and resolution, three separate VSTM tasks were used. Participants completed a change detection task with both between-class and within-class change so as to evaluate both number and resolution respectively (Awh, et al., 2007). Additionally both a color and a spatial version of the

short-term recall task (Zhang & Luck, 2008) were completed and Pmem and SD measures were extracted again to measure number and resolution respectively.

Method

Participants

Sixty-nine naïve volunteers (ages 18-32; 31 women) were recruited from the Georgia Institute of Technology community via Experimentix. All participants were right handed. Forty-five of the participants completed 10 sessions over the course of one month. Session 1 and 10 lasted approximately two hours and sessions 2-9 lasted between 40 and 60 minutes. The remaining 24 participants completed two sessions which were identical to sessions 1 and 10 for the other participants. For their participation, participants received either pay (\$10/hour) or course credit (1 credit/hour) in partial fulfillment of a course requirement. Participants who completed the eight training sessions were also paid a monetary bonus based on task performance (up to \$10 total).

Groups

Participants were randomly assigned to three groups: No-Contact Control (NCC) group, Verbal Training (VT) group, and Spatial Training (ST) group. All participants completed a test battery session on the first and last days of participation. The tasks are outlined below. Battery sessions were completed 14-33 days apart and groups were matched for intersession duration. In addition, both the VT and ST groups completed the eight intervening training sessions. Training sessions included 40-60 minutes of the adaptive verbal *N*-back and adaptive spatial *N*-back task respectively (described in detail below).

Training Session Tasks

The VT group completed eight sessions of the adaptive verbal *N*-back task. This is a continuous performance task in which participants must monitor a string of centrally presented letters. Throughout the experiment the outline of a white square was centrally presented on a black background. On each trial a capital letter printed in white appeared inside the square. After 500*ms* the letter disappeared and participants had 2500*ms* to make a response before the next trial began. On each trial, participants were asked to decide if the letter presented on the current trial matched the letter that appeared *n* trials ago. This task was adaptive in that the difficulty level changed based on participant performance. If the participant made fewer than three errors in a given block of trials, *n* increased by one on the subsequent block. Similarly if the participant made greater than five errors, *n* decreases by one; and if the participant made between three and five errors, *n* remained the same. Also, given the importance of variability in the training environment (cf. Schmidt & Bjork, 1992), if the participant completed five blocks in a row at the same level of *n*, then *n* increased by one on the following block. Each block was comprised of 20 + *n* trials and only the last 20 trials were scored (because the first *n* trials were necessarily “no match” trials). There were 30 blocks per training session for a total of 4,800 scored trials across eight days of training. Accuracy and reaction time feedback was provided at the end of each block and participants were verbally encouraged by the experimenter every 4-5 blocks.

The ST group completed eight sessions of the adaptive spatial *N*-back task. This task was conceptually identical to the adaptive verbal *N*-back task, except that the stimuli were spatial locations instead of letters. On each trial a 5 x 5 grid (white on a black

background) were presented and one of the cells were filled in red. After 500ms the filled cell disappeared and participants had 2500ms to make a response before the next trial began. On each trial, participants were asked to decide if the spatial location indicated on the current trial matches the spatial location that appeared n trials ago. Again, this task was adaptive in that the difficulty level changed based on participant performance. The adaptive schedule was identical to that used in the adaptive verbal N -back task. Each block was comprised of $20 + n$ trials, and again only the last 20 trials were scored. There were 30 blocks per training session for a total of 4,800 scored trials across eight days of training. Accuracy and reaction time feedback was provided at the end of each block and participants were verbally encouraged by the experimenter every 4-5 blocks.

For both training groups, during the first training session participants were given extensive task instructions and completed three practice blocks (1-back, 2-back, and 3-back) which were identical to the experimental blocks except that a tone sounded when an error was made.

Battery Session Tasks

For all participants, a battery of computerized tasks was administered during the first and last experimental sessions. All tests were presented using a Dell Dimensions PC computer and 24" CRT monitor using Eprime (Schneider, Eschman, & Zuccolotto, 2002) or MATLAB. The battery tasks included three tests of VSTM (change detection, color short-term recall, spatial short-term recall), two tests of WM (automated operation span and automated symmetry span tasks), two tests of general Gf (i.e., RAPM and CCF tasks), and two tests of AC (flanker and antisaccade tasks).

Visual Short-Term Memory Measures

Change Detection Task

In the change detection task (timing based on Awh, Barton, & Vogel, 2007; stimuli from Fukuda, Vogel, Mayr, & Awh, 2010), on each trial participants were presented with an arrow just above a fixation cross (200ms) cuing them to one side of the display; participants were instructed to only focus on the cued side of the display for the duration of the trial. The fixation cross remained on the screen throughout the trial. After a brief delay (200ms), a memory set appeared (500ms). The memory set consisted 2, 4, 6, 8, 12, or 16 stimuli divided evenly between the left and right sides of the display (resulting in memory set sizes of 1, 2, 3, 4, 6, and 8). There were two classes of stimuli: Ovals and rectangles (Figure 1). After the memory set disappeared, there was a retention interval (1,000ms) which included only the fixation cross followed by a test display (2000ms). The test display consisted of a single stimulus presented on each side of the display. Each stimulus was in the exact location as one of the items from the memory set. Participants were instructed to focus only on the item appearing on the cued side of the display. This target stimulus was either the same or different from the item that was in that same location during the memory set. Participants were instructed to make a button push response indicating whether or not a change occurred. There were two types of changes, between-class (rectangle to oval or oval to rectangle) and within-class (one rectangle to a different rectangle or one oval to a different oval). Participants completed 20 practice trials (5 within-class, 5 between-class, and 10 no change) followed by three experimental blocks each with 48 trials (12 within-class, 12 between-class, and 24 no

change). At the end of each block, participants were shown both their mean speed and accuracy on that block. This is a measure of VSTM capacity.

Short-Term Recall Tasks

The color and spatial short-term recall tasks used were conceptually identical except for the type of stimuli used (Zhang & Luck, 2008). On each trial in the color version of the task, six colored circles appeared at random locations around a centrally presented fixation point (100ms). After a delay (900ms), a color wheel consisting of 180 evenly distributed colors appeared. Inside the color wheel, the locations of six previously presented circles were indicated with dotted white outlines. One of the locations was cued with a solid white outline. Participants used the mouse to indicate which color on the color wheel matched the color of the circle that previously appeared at the cued location. Participants completed five practice trials followed by four experimental blocks with 60 trials each. On each trial of the spatial version of the task, six capital letters (A, B, C, D, E, F, or G) appeared at random locations around a fixation point (100ms). After a delay (900ms), a gray wheel (same dimensions as the color wheel from the color version of the experiment) appeared and one of the previously presented letters appeared in the center. Participants used the mouse to click on the gray wheel indicating at which location that letter had previously appeared. Again, participants completed five practice trials followed by four experimental blocks with 60 trials each. These tasks measure VSTM capacity.

Working Memory Measures

Automated Operation Span Task

In the automated operation span task (Unsworth, Heitz, Schrock, & Engle, 2005), participants were asked to remember a series of 3-7 letters in order. Letter presentations

(1000ms each) were interleaved with simple to-be-solved math problems. Time limit for math subtask were individualized and determined by practice phase performance. After all letters were presented, participants recalled all letters in the order in which they were presented. There were a total of 75 trials. Prior to the task, participants completed 10 trials of the letter task alone, 15 trials of the math task alone, and 6 trials of the combined task. Feedback was provided after each trial. This is a measure of verbal WM capacity.

Automated Symmetry Span Task

In the automated symmetry span task (Unsworth, et al., 2005), participants were asked to remember a series of 2-5 spatial locations. On each trial a 4x4 grid was presented in which one of the 16 possible locations was filled in red (650ms each). Participants were asked to remember the location of the red square. Between each location presentation, participants were presented with a geometric figure and asked to determine whether or not the figure was symmetrical about the vertical axis. Time limit for the symmetry judgment sub task were individualized and determined by practice phase performance. After all spatial locations have been presented, participants were asked to reproduce all spatial locations in the order in which they were presented. There were a total of 42 trials. Prior to the task, participants completed 10 trials of the locations task alone, 15 trials of the symmetry task alone, and 6 trials of the combined task. Feedback was provided after each trial. This is a measure of spatial WM capacity.

General Fluid Intelligence Measures

Raven's Advanced Progressive Matrices Task

In the Raven's advanced progressive matrices task (RAPM; Raven, 1990), on each trial participants were presented with a 3 x 3 matrix. Eight of the cells were filled

with related line drawings. The lower right cell was blank and participants had to decide which of eight possible line drawings best fit into that cell. Problems got progressively harder throughout the task. The task ended after either 18 problems had been completed or 20 minutes had elapsed. Participants completed odd numbered trials during one session and even numbered trials during the other session, order was counterbalanced across participants. This is a measure of *Gf*.

Cattell's Culture Fair Task

In the CCF task, participants completed four subtasks (series completion, odd elements, matrix completion, and dot task; Cattell, 1949). In the series completion task (7 problems per session), participants saw three simple line drawings that together created a pattern. Participants had to decide which of six possible similar pictures best completes the pattern. In the odd elements task (7 problems per session), participants saw five simple line drawings and had to determine which two drawings did not belong with the rest. In the matrix completion task (7 problems per session), participants were presented with either a 2x2 or 3x3 matrix. One of the cells was empty and participants had to decide which of four possible alternatives best fit into the empty cell. Occasionally the matrices were partially obscured by “cut outs” or missing information. In the dot task (6 problems per session), participants were presented with a simple line drawing with a dot present. Participants had to then determine in which of five possible alternative drawings (without dots) would allow for a dot to be placed in a comparable position to the sample drawing. Participants completed odd numbered problems during one session and even numbered problems during the other session (order was counterbalanced across participants) for a total of 27 problems per session. This is also a measure of *Gf*.

Attentional Control Measures

Antisaccade Task

In the antisaccade task (Unsworth, et al., 2004), on each trial began with a blank screen (400ms) followed by a three asterisk fixation (200ms, 600ms, 1000ms, 1400ms, and 1800ms) and another blank screen (10ms). Next a cue appeared (equal sign; two blinks with 100ms on and 50ms off) either on the left or the right of the display. A target letter (B, P, or R) appeared briefly (100ms) on the opposite side of the display and was immediately masked with an H (50ms) followed by an 8 which remained on the screen until the participant responded. Participants completed 60 practice trials in which cues and targets were all centrally presented with feedback on each trial. Participants also completed 10 trials of the experimental task also with feedback. Finally, participants completed 60 experimental trials with block feedback at the end; 30 with a left cue and 30 with a right cue. The three target letters as well as the five fixation durations were evenly distributed across trials. This is a measure of AC.

Flanker Task

In the flanker task (Eriksen & Eriksen, 1974), participants were presented with a centrally presented fixation dot (200ms) followed by five arrows (e.g., >>>>>; 100ms). Participants were asked to determine whether the central arrow was facing the right or left and respond with a button push. Half of the trials were congruent (e.g., >>>>>) and half of the trials were incongruent (e.g., >><>>). There were an equal number of left and right responses. Five delays separating each trial were evenly distributed across trials (200ms, 600ms, 1000ms, 1400ms, and 1800ms). There were a total of 16 practice trials

with feedback on each trial and 120 experimental trials with block feedback at the end. This is also a measure of AC.

Analyses

All battery task data and training task data were analyzed using SPSS. Training task performance was evaluated using a Training Session (1-8) x Group (VT vs. ST) repeated measures ANOVAs on the maximum difficulty (i.e., maximum level of n) achieved. To assess the efficacy of cognitive training on untrained measures of WM (automated operation span and automated symmetry span scores) and Gf (RAPM and CCF scores), composite scores were generated. Composite scores were calculated with the method previously used by Redick et al. (2011). When appropriate, principal axis factoring was used to combine contributing raw scores and to extract a single factor score separately for BS1 and BS2. No composite score was calculated for measures of AC, because the component tasks (antisaccade and flanker) unexpectedly failed to correlate with each other. Thus, Gf and WM transfer efficacy was then determined by submitting the composite scores to separate Time (BS1 vs. BS2) x Group repeated measures ANOVAs. To assess training transfer efficacy on AC, both the antisaccade data flanker data were separately submitted to the same Time x Group repeated ANOVAs. To assess training transfer efficacy on VSTM, separate Time x Group repeated measures ANOVAs were performed on within-class and between-class accuracy scores on the change detection task. For the short-term recall tasks, maximum likelihood estimation was used to fit a von Mises distribution model (chosen because of the circular stimulus space) to each participant's data (Zhang & Luck, 2008, 2009). Briefly, this model includes two parameters: P_{mem} and SD (Zhang & Luck, 2008, 2009). The P_{mem} parameter reflects

the height of the von Mises distribution (indicating the probability that the probed item is not in memory) subtracted from 1. Pmem indicates the probability that the cued item exists in memory and is a measure of the number of items held in VSTM. The SD measure reflects the width of the von Mises distribution and indicates the resolution with which these items are held. Pmem and SD were each submitted to a Time x Group repeated measures ANOVA individually for the color and spatial short-term recall tasks. Composite scores were also calculated for both number (color Pmem, spatial Pmem, and between-class change detection accuracy) and resolution (color SD, spatial SD, and within-class change detection accuracy) as described above. Post hoc analyses were conducted via Bonferroni corrected one-tailed independent samples t-tests on the difference scores (BS2 – BS1) where appropriate at a significance cutoff of $p = .016$ unless otherwise specified. In all cases, post hoc analyses were only conducted if a significant interaction was present.

Procedure

Participants were randomly assigned to one of the three experimental groups. All participants arrived in the laboratory on the first day and read and sign the consent form prior to the start of the experiment. Each participant completed all battery tasks over the course of approximately 2.5 hours. The order of tasks was counterbalanced across participants. At the end of the first session, participants were paid or awarded credit for their participation and reminded of the date and time of their next session. Participants in the two training groups came back to the laboratory for 8 additional training sessions. During the first training session, participants read the instructions and completed three practice blocks (1-back, 2-back, and 3-back). Accuracy and reaction time feedback was

provided at the end of each practice block. After training was complete, participants engaged in 30 experimental blocks. At the end of each training session participants were again be paid or awarded credit and reminded of the time and date of their next session. All participants completed their final session 14-33 days after their first session (matched groups). This final session was identical to the first session except that the order of battery tasks was again counterbalanced across participants.

Results

Two participants (one from the NCC group and one from the ST group) failed to complete the all of the required sessions and were removed from the analysis. One additional participant from the NCC group was removed from the analysis because she failed to comply with instructions on 4 (automated operation span, automated symmetry span, color short-term recall, and spatial short-term recall) of the 9 battery tasks and performed greater than two standard deviations below the mean during session 1 on 4 (flanker, CCF, RAPM, and change detection) of the remaining 5 battery tasks.

Training Tasks

As previously noted, the first goal of cognitive training is to improve performance on the trained task (Willis, et al., 2006). This was evaluated with a Training Session x Group repeated measures ANOVA on the maximum difficulty achieved (i.e., max n) on each training day (Figure 2a). The assumption of sphericity was violated for the main effect of Session ($p < .001$), thus degrees of freedom were corrected using the Huynh-Feldt adjustment. The main effect of Session was significant, $F(3.8, 162.1) = 62.33$, $p < .001$, $\eta_p^2 = .60$, with performance improving across the 8 training sessions. The main effect of Group was also significant, $F(1, 42) = 6.01$, $p = .018$, $\eta_p^2 = .13$, with the VT

group achieving an overall higher level of n than the ST group. Finally, the Interaction approached significance, $F(3.9, 162.1) = 2.22$, $p = .072$, $\eta_p^2 = .05$, with a trend toward the VT group showing larger improvement over the course of training compared to the ST group.

When only the overall gain across training was considered (i.e., maximum level of n on training session 8 minus maximum level of n on training session 1), a two-tailed independent samples t -test revealed a significant difference between the groups, $t(42) = 2.56$, $p = .015$ (Figure 2b). These data suggest that overall, the VT group showed greater improvement from training session 1 to training session 8 (max n increased from 4.8 to 9.5) compared to the ST group (max n increased from 4.5 to 7.8).

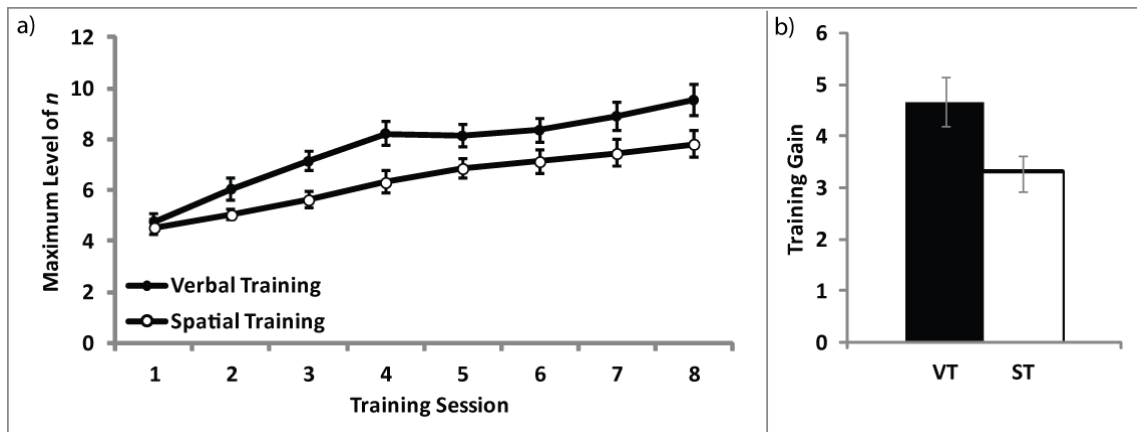


Figure 2.(a) Training Task Data (b) Training Gain

Battery Tasks

Visual Short-Term Memory

Change Detection Task

For the change detection task, the dependent variable of interest was accuracy.

One participant (NCC group) performed greater than three standard deviations below the

mean during the first battery session and was removed from the analysis. Overall accuracy was assessed with a Time x Group repeated measures ANOVA (Figure 3). Both the main effect of Time, $F(1,62) = 11.01, p = .002, \eta_p^2 = .15$, and the Interaction, $F(2,62) = 6.69, p = .002, \eta_p^2 = .18$, were significant. The main effect of Group, $F(2,62) = 1.77, p = .179, \eta_p^2 = .05$, was not significant. Post hoc analysis evaluating the efficacy of training revealed a significant difference between the VT and NCC groups, $t(41) = 2.56, p = .014$, and the ST and NCC groups, $t(41) = 3.18, p = .003$. There was no difference between the VT and ST groups, $t(42) = -.75, p = .461$.

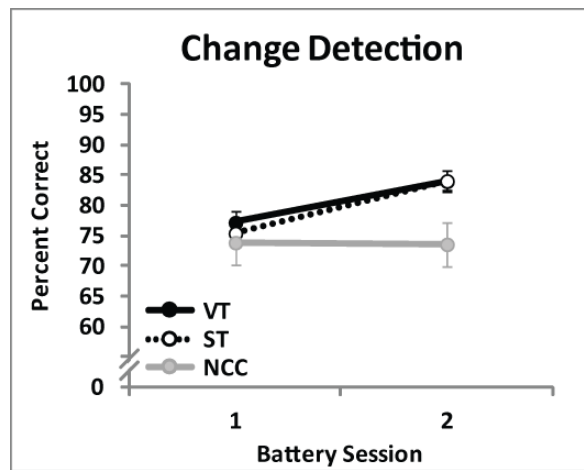


Figure 3. Change Detection Task Data

Next the influence of WM training on the number and resolution subprocesses of VSTM were evaluated individually. Training-related improvements were evaluated using a Time x Group repeated measures ANOVA on between-class change trial accuracy (Figure 4a). Again, both the main effect of Time, $F(1,62) = 12.25, p = .001, \eta_p^2 = .17$, and the Interaction, $F(2,62) = 10.18, p < .001, \eta_p^2 = .25$, were significant. The main effect of Group approached significance, $F(2,62) = 2.99, p = .058, \eta_p^2 = .09$. Post hoc

evaluation of the efficacy of training revealed a significant difference between the VT and NCC groups, $t(41) = 3.02, p = .002$, and the ST and NCC groups, $t(41) = 3.84, p < .001$. There was no difference between the VT and ST groups, $t(42) = -1.27, p = .210$. Similarly, training-related improvements on resolution was evaluated using a Time x Group ANOVA on within-class change trial accuracy (Figure 4b). The main effect of Time, $F(1,62) = 13.47, p = .001, \eta_p^2 = .18$, the main effect of Group, $F(2,62) = 3.70, p = .030, \eta_p^2 = .11$, and the Interaction, $F(2,62) = 7.05, p = .002, \eta_p^2 = .18$, were all significant. Again, post hoc evaluation of the efficacy of training revealed a significant difference between the VT and NCC groups, $t(41) = 2.68, p = .006$, and the ST and NCC groups, $t(41) = 3.08, p = .002$, but no significant difference between the VT and ST groups, $t(41) = -.47, p = .642$.

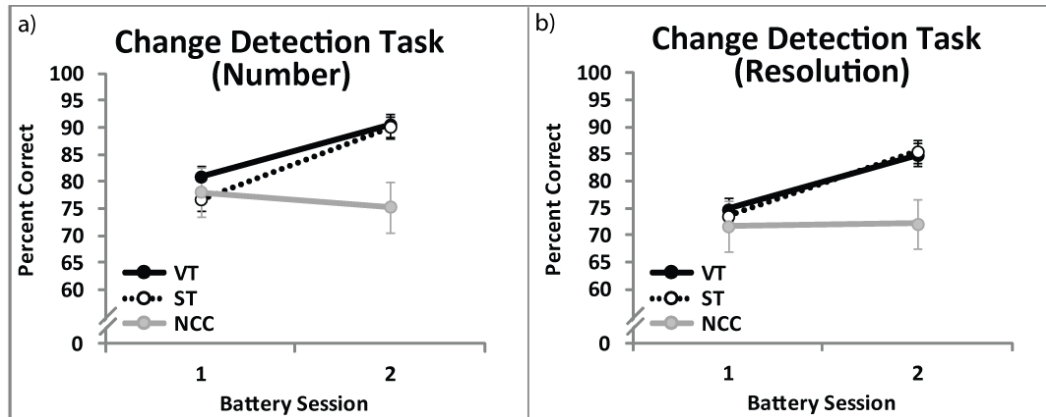


Figure 4. (a) Change Detection Task Measures of Number (b) Change Detection Task Measures of Resolution

Capacity estimates at each set size were calculated using Cowan's k ($k = [\text{set size} * (\text{hit rate} + \text{correct rejection rate} - 1)]$; Cowan, 2001) for both BS1 and BS2. Separate Time (BS1 vs. BS2) x Set Size (1, 2, 3, 4, 6, and 8) repeated measures ANOVAs were

conducted for each of the three groups. For the VT group (Figure 5a), the main effect of Time, $F(1,21) = 6.42, p = .019, \eta_p^2 = .23$, and the main effect of Set Size, $F(1.9,39.1) = 11.90, p < .001, \eta_p^2 = .36$, were significant and the Interaction, $F(2.0,41.4) = 2.29, p = .115, \eta_p^2 = .1$, was not significant. For the ST group (Figure 5b), the main effect of Time, $F(1,21) = 21.86, p < .001, \eta_p^2 = .51$, the main effect of Set Size, $F(1.9,40) = 11.20, p < .001, \eta_p^2 = .35$, and the Interaction, $F(2.9,61.2) = 4.57, p = .006, \eta_p^2 = .18$, were all significant. Finally, for the NCC group (Figure 5c), the main effect of Set Size, $F(2.2,44.3) = 4.46, p = .014, \eta_p^2 = .182$, was significant, but neither the main effect of Time, $F(1,20) = .28, p = .603, \eta_p^2 = .01$, nor the Interaction, $F(2.2,43.1) = .38, p = .698, \eta_p^2 = .12$, was significant.

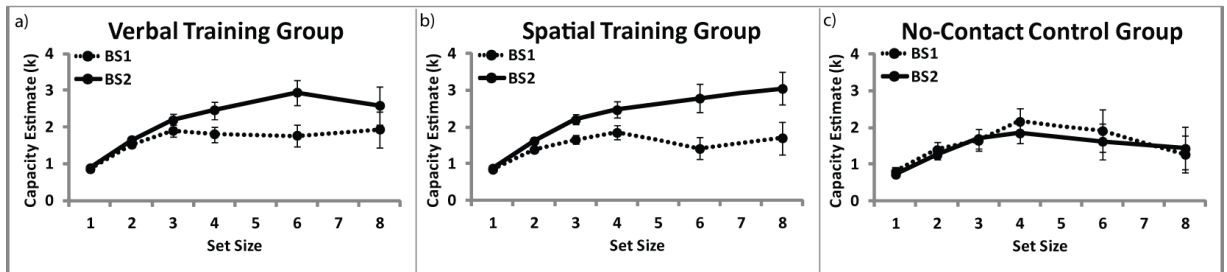


Figure 5. (a) Verbal Training Group Visual Short-Term Memory Capacity Estimates (b) Spatial Training Group Visual Short-Term Memory Capacity Estimates (c) No-Contact Control Group Visual Short-Term Memory Capacity Estimates

Short-Term Recall Tasks

For both the short-term recall tasks, training-related performance improvement was measured by extracting Pmem and SD scores and comparing these scores from BS1 to BS2. Three participants (two from the VT group and one from the ST group) were removed from the analysis because they failed to follow instructions on one or both of the tasks. For the color short-term recall task, training efficacy on Pmem was assessed with a

Time x Group repeated measures ANOVA (Figure 6a). Neither the main effect of Time, $F(1,62) = 2.39, p = .128, \eta_p^2 = .04$, the main effect of Group, $F(2,62) = .76, p = .473, \eta_p^2 = .02$, nor the Interaction, $F(2,62) = .49, p = .617, \eta_p^2 = .02$, was significant. Training efficacy on SD was also assessed with a Time x Group repeated measures ANOVA (Figure 6b). Again, neither the main effect of Time, $F(1,62) = .24, p = .626, \eta_p^2 = .004$, the main effect of Group, $F(2,62) = .25, p = .782, \eta_p^2 = .008$, nor the Interaction, $F(2,62) = .83, p = .439, \eta_p^2 = .03$, was significant indicating no effect of training on either measure.

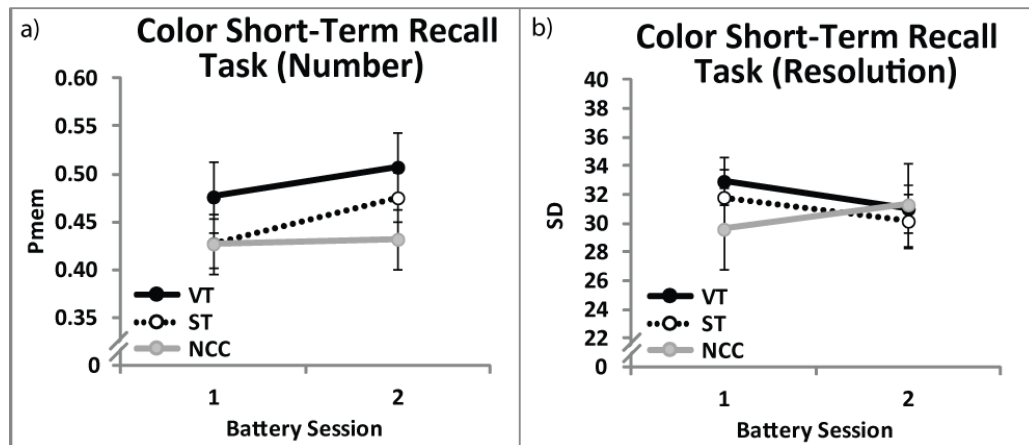


Figure 6. (a) Color Short-Term Memory Task Measures of Number (b) Color Short-Term Memory Task Measures of Resolution

For the spatial short-term recall task, training efficacy on Pmem and SD were again assessed individually using a Time x Group repeated measures ANOVA (Figure 7a). For the Pmem measure, neither the main effect of Time, $F(1,61) = .76, p = .388, \eta_p^2 = .01$, the main effect of Group, $F(2,61) = 2.60, p = .082, \eta_p^2 = .08$, nor the Interaction, $F(2,61) = .87, p = .422, \eta_p^2 = .03$, was significant. For the SD measure (note that a smaller SD measure indicates better performance; Figure 7b), both the main effect of

Time, $F(1,61) = 4.39$, $p = .040$, $\eta_p^2 = .07$, and the Interaction, $F(2,61) = 6.31$, $p = .003$, $\eta_p^2 = .17$, were significant. The main effect of Group, $F(2,61) = 1.31$, $p = .277$, $\eta_p^2 = .04$, was not significant. Post hoc evaluation revealed a significant difference between the VT and ST groups, $t(40) = 3.20$, $p = .004$, as well as between the VT and NCC groups, $t(40) = 3.29$, $p = .001$. The difference between ST and NCC groups was not significant, $t(40) = .40$, $p = .346$. These data indicate that both the ST and NCC groups showed similar improvement from BS1 to BS2 and the VT group performed worse on BS2 compared to BS1.

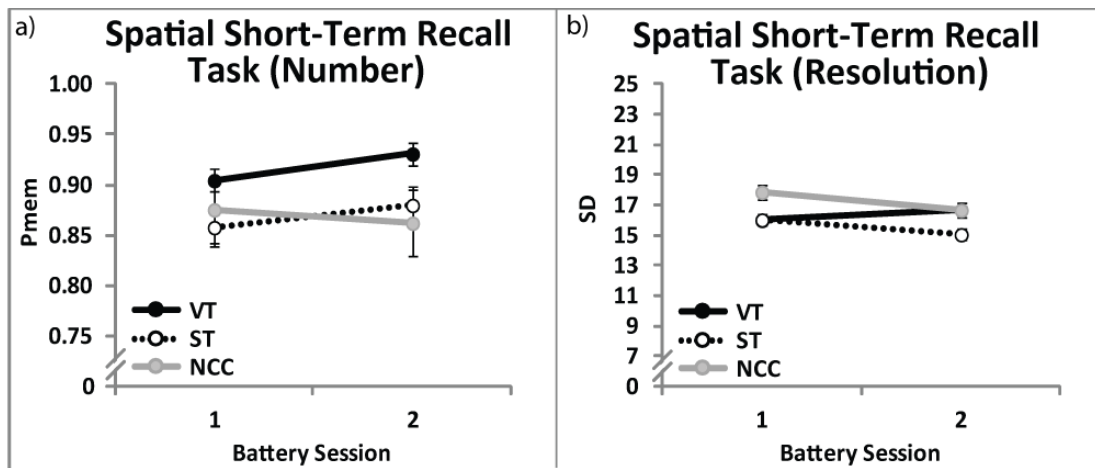


Figure 7. (a) Spatial Short-Term Memory Task Measures of Number (b) Spatial Short-Term Memory Task Measures of Resolution

Capacity estimates (K) were calculated by multiplying the set size (i.e., 6) by the probability that a probed item was present in memory for each individual both during BS1 and BS2 for both the color and spatial short-term tasks (Zhang & Luck, 2008, 2011). These capacity estimates were submitted by a Time x Group repeated measures ANOVA. For the color short-term recall task, neither the main effect of Time, $F(1,62) = 2.39$, $p = .127$, $\eta_p^2 = .04$, the main effect of Group, $F(2,62) = .76$, $p = .473$, $\eta_p^2 = .02$, nor the

Interaction, $F(2,62) = .49$, $p = .618$, $\eta_p^2 = .02$, was significant. For the spatial short-term recall task, again neither the main effect of Time, $F(1,61) = .75$, $p = .389$, $\eta_p^2 = .01$, the main effect of Group, $F(2,61) = 2.60$, $p = .083$, $\eta_p^2 = .08$, nor the Interaction, $F(2,61) = .87$, $p = .422$, $\eta_p^2 = .03$, was significant. These data indicate that the number of items held in VSTM as measured on the short-term recall task did not significantly change with training.

These results can be contrasted with similar analyses on the change detection data at set size 6 (Figure 8) indicating both a significant main effect of Time, $F(1,62) = 10.47$, $p = .002$, $\eta_p^2 = .14$, and significant Interaction, $F(2,62) = 5.17$, $p = .008$, $\eta_p^2 = .14$. The main effect of Group, $F(2,62) = .682$, $p = .509$, $\eta_p^2 = .02$, was not significant in this analysis. Post hoc evaluation revealed a significant difference between the VT and NCC groups, $t(41) = 2.74$, $p = .005$, and between the ST and NCC groups, $t(41) = 2.94$, $p = .003$. The difference between VT and ST groups, $t(41) = -.319$, $p = .952$, was not significant.

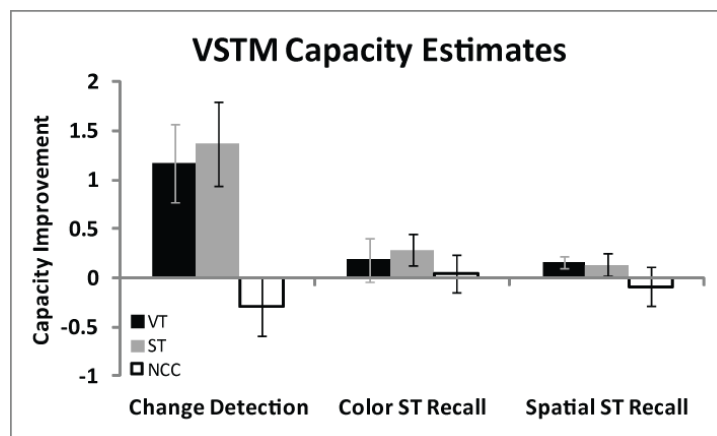


Figure 8. Visual Short-Term Memory Capacity Estimates across Tasks

VSTM Composites

All measures of VSTM number were significantly with each other: Change detection between-class change accuracy correlated with spatial short-term recall Pmem scores ($r = .28, p = .024$); change detection between-class change accuracy correlated with color short-term recall Pmem scores ($r = .35, p = .004$); and spatial short-term recall Pmem scores correlated with color short-term recall Pmem scores ($r = .50, p < .001$). Therefore, to achieve a coherent general measure of VSTM number, factor scores were extracted from the measures of number from the change detection task and each of the short-term recall tasks during BS1 and BS2 to create a VSTM number composite score (Figure 9a). These scores were then submitted to a Time x Group repeated measures ANOVA. Neither the main effect of Time, $F(1,60) = .006, p = .941, \eta_p^2 = 0$, nor the main effect of Group, $F(2,60) = 2.18, p = .122, \eta_p^2 = .07$, were significant. The Interaction, $F(2,60) = 4.66, p = .013, \eta_p^2 = .13$, was significant. Post hoc analyses revealed that the VT and ST groups, $t(39) = -1.51, p = .280$, did not significantly differ, nor did the VT and NCC groups, $t(40) = 1.74, p = .089$. There was a significant difference between the ST and NCC groups, $t(41) = 2.64, p = .012$.

All measures of VSTM resolution also significantly correlated with each other, however, the pattern of correlations was unexpected. Change detection within-class change accuracy and spatial short-term recall SD score were negatively correlated ($r = -.35, p = .005$); this was expected given that both high accuracy and a low SD scores indicate good resolution performance. Spatial short-term recall SD scores also negatively correlated with color short-term recall scores ($r = -.35, p = .005$), while change detection within-class change accuracy and color short-term recall SD scores positively correlated

($r = .33, p = .007$); the direction of these effects is surprising and difficult to interpret. It is unclear why good performance on the color short-term recall task would correlate with bad performance on the spatial short-term memory test or the change detection task measures of resolution. This may suggest that an inconsistency between “resolution processes” as measured by the various tasks. Thus, composite scores should be interpreted with caution. Still, factor scores were extracted from the measures of number from the change detection task and each of the short-term recall tasks during BS1 and BS2 to create a VSTM resolution composite score (Figure 9b). Composite scores were then submitted to a Time x Group repeated measures ANOVA. Neither the main effect of Time, $F(1,60) = .002, p = .964, \eta_p^2 = 0$, the main effect of Group, $F(2,60) = .96, p = .333, \eta_p^2 = .03$, nor the Interaction, $F(2,60) = 1.17, p = .318, \eta_p^2 = .04$, was significant.

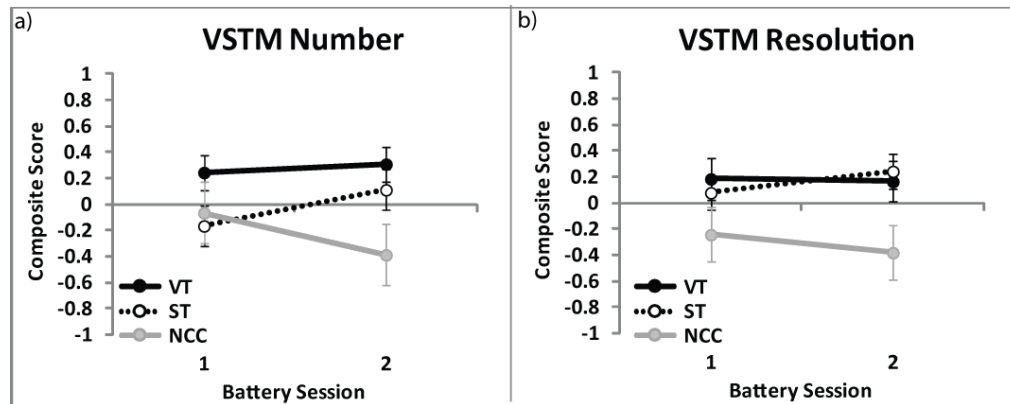


Figure 9. (a) Visual Short-Term Memory Composite Scores of Number (b) Visual Short-Term Memory Composite Scores of Resolution

Working Memory

Automated Span Tasks

For both the automated operation span and automated symmetry span tasks, training-related performance improvement was measured by comparing the total score of

correct items from BS1 to the total score of correct trials from BS2. Two participants were removed from the automated operation span task because one (NCC group) achieved less than 85% accuracy on the math tasks (which was a task requirement) and the other (ST group) recalled zero items correctly during the first battery session which was greater than three standard deviations below the mean. Training-related improvements were evaluated with a Time x Group repeated measures ANOVA. For the automated operation span task (Figure 10a), neither the main effect of Time, $F(1,62) = 1.43, p = .246, \eta_p^2 = .02$, nor the Interaction, $F(2,62) = 1.95, p = .151, \eta_p^2 = .06$, were significant. The main effect of Group, $F(2,62) = 3.33, p = .042, \eta_p^2 = .10$, was significant with the VT group achieving higher scores than the ST group ($p = .037$). For the automated symmetry span tasks (Figure 10b), the main effect of Time was significant, $F(2,63) = 7.90, p = .007, \eta_p^2 = .11$, with participants achieving higher scores during the second battery session compared to the first. Neither the main effect of Group, $F(2,63) = 1.61, p = .209, \eta_p^2 = .05$, nor the Interaction, $F(2,63) = 2.34, p = .105, \eta_p^2 = .07$, was significant.

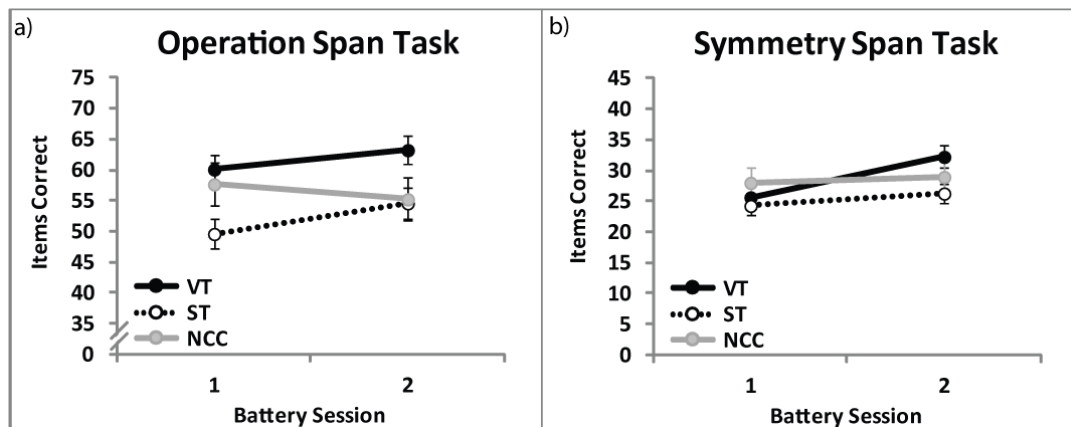


Figure 10. (a) Operation Span Task Data (b) Symmetry Span Task Data

Working Memory Composite

Automated operation span and automated symmetry span scores were significantly correlated ($r = .30, p = .016$), thus to achieve a coherent general measure of WM, factor scores were extracted from both the automated operation span and symmetry tasks during BS1 and BS2 to create a WM composite score. These WM composite scores were then submitted to a Time x Group ANOVA (Figure 11). Neither the main effect of Time, $F(1,61) = .02, p = .892, \eta_p^2 = 0$, nor the main effect of Group, $F(2,61) = 2.8, p = .069, \eta_p^2 = .10$, were significant. The Interaction, $F(2,61) = 3.88, p = .026, \eta_p^2 = .113$, was significant. Post hoc analysis revealed that the VT group showed a significantly larger training effect than the NCC group, $t(41) = 2.53, p = .008$. Neither the difference between the VT and ST groups, $t(41) = 1.96, p = .058$, nor the ST and NCC groups, $t(40) = 1.03, p = .155$, were significant. These data indicate that only the VT group showed training-related transfer to measures of WM.

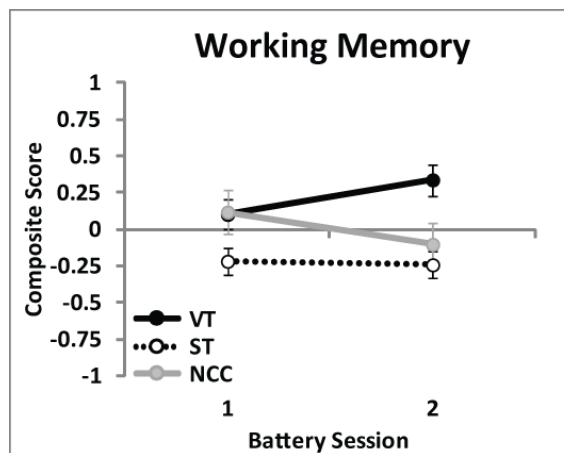


Figure 11. Working Memory Task Composite Scores

General Fluid Intelligence

Ravens Advanced Progressive Matrices and Cattell's Culture Fair Tasks

Training-related performance improvement on *Gf* was measured by comparing the total number of correct items from BS1 to the total number of correct items from BS2 for both the RAPM task and the CCF task. For the RAPM task, a Time x Group repeated measures ANOVA was conducted (Figure 12a). Neither the main effect of Time, $F(1,63) = .08, p = .776, \eta_p^2 = .001$, the main effect of Group, $F(2,63) = .09, p = .917, \eta_p^2 = .003$, nor the Interaction, $F(2,63) = 1.68, p = .195, \eta_p^2 = .05$, was significant. A Time x Group repeated measures ANOVA was similarly conducted for the CCF task (Figure 12b). Again, neither the main effect of Time, $F(1,63) = .26, p = .615, \eta_p^2 = .004$, the main effect of Group, $F(2,63) = 1.85, p = .165, \eta_p^2 = .06$, nor the Interaction, $F(2,63) = .65, p = .528, \eta_p^2 = .02$, was significant.

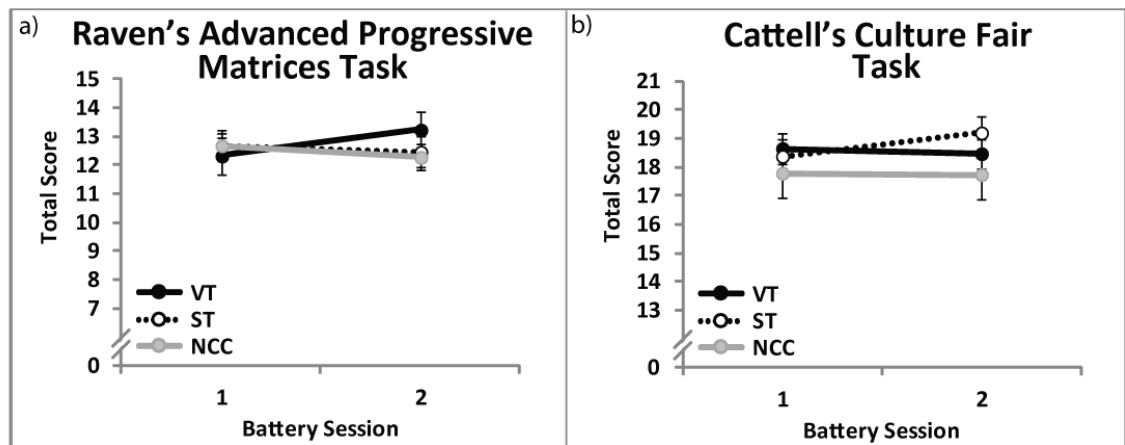


Figure 12. (a) Raven's Advanced Progressive matrices Task Data (c) Cattell's Culture Fair Task Data

General Fluid Intelligence Composite

RAPM and CCF scores were significantly correlated ($r = .32, p = .008$), therefore to compute a general *Gf* composite score factor scores were extracted from both the RAPM and CCF tasks during BS1 and BS2 separately. A Time x Group repeated measures ANOVA was then performed on these *Gf* composite scores (Figure 13). Neither the main effect of Time, $F(1,63) = 0, p = 1.000, \eta_p^2 = 0$, the main effect of Group, $F(2,63) = .69, p = .504, \eta_p^2 = .02$, nor the Interaction, $F(2,63) = .36, p = .702, \eta_p^2 = .01$, was significant.

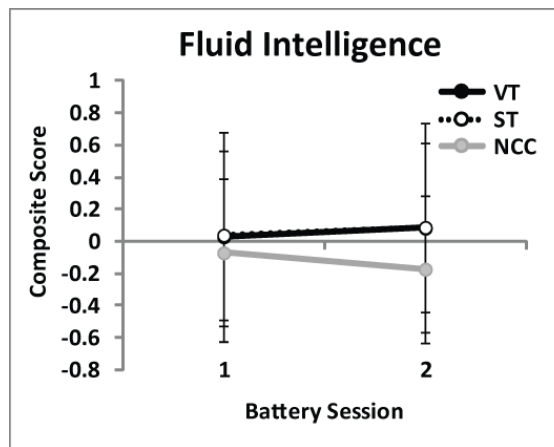


Figure 13. Fluid Intelligence Task Composite Scores

Attentional Control

Flanker and Antisaccade Tasks

Training-related performance improvement on AC processes was evaluated by comparing BS1 performance to BS2 performance on both the flanker task and the antisaccade task. For the flanker task, two participants were removed from the analysis (one from the VT group and one from the ST group) due to accuracy greater than 2.5

standard deviations below the mean on BS1. RTs on congruent trials were subtracted from RTs on incongruent trials to obtain a difference score representing the amount of interference between the two conditions. These difference scores were submitted to a Time x Group repeated measures ANOVA (Figure 14a). The main effect of Time, $F(1,61) = 15.61, p < .001, \eta_p^2 = .20$, was significant with participants showing less interference after training. The Interaction, $F(2,61) = 2.64, p = .080, \eta_p^2 = .08$, approached significance. The main effect of Group, $F(2,61) = .44, p = .645, \eta_p^2 = .01$, was significant. For the antisaccade task, RTs for correct trials were submitted to a Time x Group repeated measures ANOVA (Figure 14b). The main effect of Time, $F(1,63) = 25.24, p < .001, \eta_p^2 = .29$, was significant and the Interaction, $F(2,63) = 2.91, p = .062, \eta_p^2 = .09$, approached significant. The main effect of Group, $F(2,63) = .34, p = .712, \eta_p^2 = .01$, was not significant.

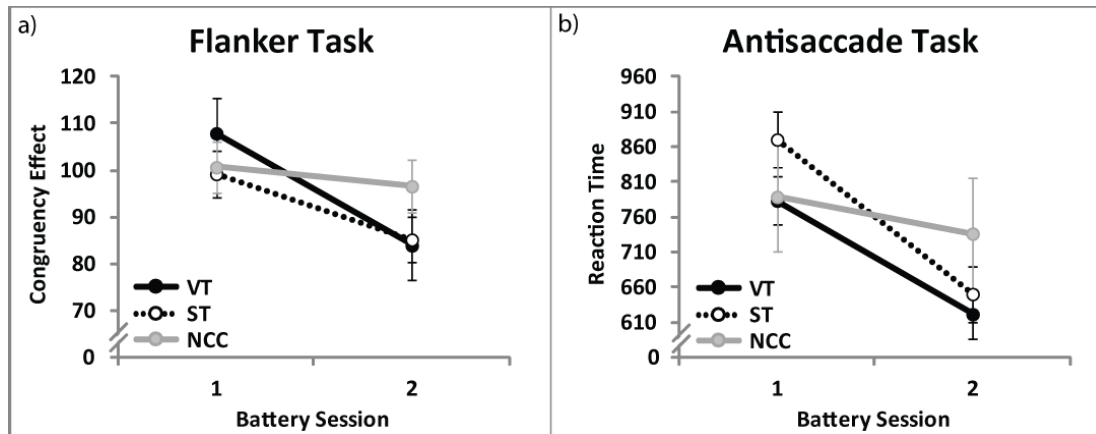


Figure 14. (a) Flanker Task Data (b) Antisaccade Task Data

Visual inspection of the data as well as post hoc analyses comparing the VT and ST groups reveal similar improvement from BS1 to BS2 for both the flanker, $t(40) = -1.08, p = .285$, and the antisaccade, $t(42) = 1.10, p = .277$, tasks. To further investigate this trend, data were collapsed across training group and data were resubmitted to a Time

x Group (VT/ST vs. NCC) repeated measures ANOVA independently for each task. For the flanker task, there was a significant main effect of Time, $F(1,62) = 9.49, p = .003, \eta_p^2 = .13$, and the Interaction, $F(2,62) = 3.97, p = .051, \eta_p^2 = .06$, approached significance. The main effect of Group, $F(2,62) = .39, p = .535, \eta_p^2 = .01$, was not significant. Post hoc analysis revealed that the VT/ST group showed a significant training effect compared to the NCC group, $t(62) = -1.99, p = .026$. For the antisaccade task, the main effect of Time, $F(1,64) = 15.93, p < .001, \eta_p^2 = .20$, and the Interaction, $F(2,64) = 5.16, p = .026, \eta_p^2 = .08$, were both significant. The main effect of Group, $F(1,64) = .192, p = .663, \eta_p^2 = .003$, was not significant. Again, post hoc analysis revealed that the VT/ST group showed a significant training effect compared to the NCC group, $t(64) = -2.27, p = .013$.

Attentional Control Composite

Flanker and Antisaccade performance was not significantly correlated ($r = .13, p = .317$) and therefore composite scores were not calculated.

CHAPTER 6

GENERAL DISCUSSION

The purpose of the current study was twofold. The first goal was to develop an optimized training design to promote high-level performance on the training task as well as transfer to untrained cognitive skills. The second goal was to investigate the efficacy of WM training on a variety of cognitive functions with a particular focus on VSTM (a previously underinvestigated process in the cognitive training literature). The data provide unique insights into each of these aims to help paint a coherent picture of cognitive training efficacy more generally.

The Training Task

The goal of applying recommendations from the skill and cognitive training literatures was to create an optimized training design to promote learning. One recommendation that stands out in both of these literatures is the importance of variability during training (c.f. Morrison & Chein, 2011; Schmidt & Bjork, 1992). For this reason, careful considerations were made in order to ensure a continuously variable training environment. Many training studies have used an adaptive design (e.g., Jaeggi, et al., 2008; Jaeggi, et al., 2010; Redick, et al., 2011; Schwarb, et al., under review) allowing participant performance to adjust task difficulty. The current study was unique in that it made the additional requirement that no individual could complete more than five blocks in a row at the same level of difficulty ensuring that once the participant reached a plateau, training difficulty would continue to vary. This design was affective in that every participant showed improvement (and in some cases, large improvements)

across the eight training tasks. For the VT group, by training session eight participants were able to perform at levels of n that were 3-12 (mean = 4.7 ± 2.2) higher than during session one. Similarly, for the VS group, by training session eight participants were able to perform at levels of n that were 1-8 (mean = 3.3 ± 1.6) higher than during session one. Thus it appears that the optimized design indeed facilitated training improvement across eight sessions.

While the current study does not allow for a direct comparison between the optimized training design and a sub-optimal design, we have previously completed a cognitive training study that was similar in many ways to the current study, without the mandatory variability component (Schwarb, et al., under review) which may be interesting to consider. In that study, participants from the Georgia Institute of Technology completed both a spatial and a verbal version of the n -back task which was also adaptive in that difficulty was adjusted based on participant performance, however, if participants continued to perform moderately well (i.e., between three and five errors per block), they could remain at the same level of difficulty until the end of the session. As in the current experiment, quantitative feedback was provided after every block and participants were verbally encouraged throughout and given a monetary bonus for improved performance. Acknowledging that this is not an optimal comparison and that participants in that study completed the experiment between 8 and 14 months before the participants in the current study, Training Session (1-8) x Experiment (Schwarb et al., under review vs. current study) repeated measures ANOVAs were conducted for the verbal task and the spatial task separately. Because participants in the previous study (Schwarb, et al., under review) only completed 18 blocks of trials per task in any given

training session, the current data were reanalyzed to only include the first 18 blocks. For the verbal task (Figure 15a), the assumption of sphericity was violated and data were corrected using the Huynh-Feldt adjustment. There was a significant main effect of Training Session, $F(6.2,291.5) = 53.24, p < .000, \eta_p^2 = .53$, Group, $F(1,47) = 16.36, p < .001, \eta_p^2 = .26$, and Interaction, $F(6.2,291.5) = 5.58, p < .001, \eta_p^2 = .11$. For the spatial task (Figure 15b), the assumption of sphericity was again violated and data were corrected using the Huynh-Feldt adjustment. Both the main effect of Training Session, $F(5.2,244.9) = 48.5, p < .001, \eta_p^2 = .508$, and Interaction, $F(5.2,244.9) = 2.36, p = .038, \eta_p^2 = .05$, were significant. The main effect of Group, $F(1,47) = 3.84, p = .06, \eta_p^2 = .08$, approached significance. Thus, for both the spatial and the verbal adaptive n -back task, participants in the current study demonstrated greater improvement across training than the participants in the previous study. Thus while further research is necessary to confirm this finding, these data suggest that consistently varying training task difficulty could promote better learning and performance during training.

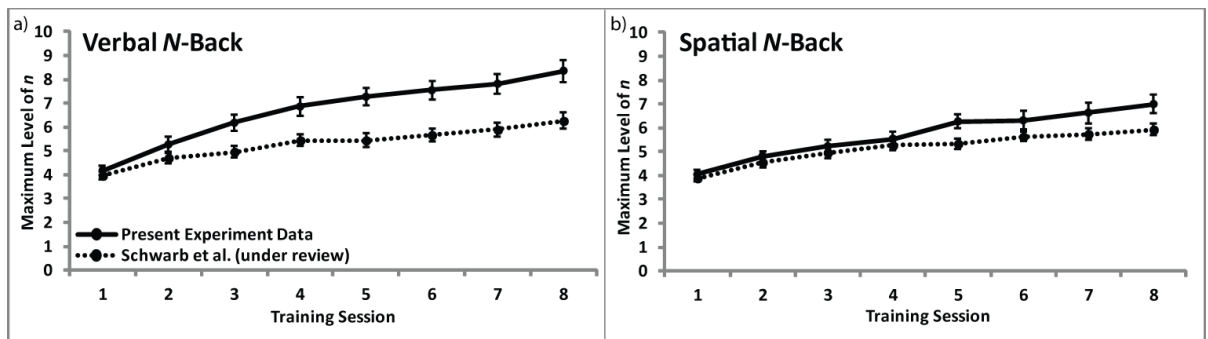


Figure 15. (a) Optimized Verbal Training Task Comparison Data (b) Optimized Spatial Training Task Comparison Data

Transfer of Cognitive Skill

Working Memory

Data regarding WM training-related improvements on other measures of WM are mixed in the literature. For example, several studies report no effect of training on a variety of measures of WM including the reading span task (Jaeggi, et al., 2008; Schmiedek, Lövdén, & Lindenberger, 2010), rotation span task (Li, et al., 2008; Schmiedek, et al., 2010), symmetry span task (Redick, et al., 2011; Schwarb, et al., under review), and the running letter span task (Redick, et al., 2011). Other studies, however, do report significant training-related improvements on measures of WM including the operation span task (Schwarb, et al., under review) and a composite temporary measure which included both the operation span and symmetry span tasks (Chein & Morrison, 2010).

Our data provide additional evidence for training-related improvements on measures of WM. While both the automated operation span and automated symmetry span data show a trend toward a significant Group x Time interaction, this effect was not significant. However, when this data was compiled into a composite measure, the analysis had sufficient power to show a significant effect. Interestingly, training-related improvements were only evident when participants were trained with a verbal adaptive *n*-back task. At first glance this is somewhat surprising. If what *n*-back training does is enhance participant's WM capacity, we might expect to see transfer to other measures of WM for both the VT and the ST groups. However, it is also important to remember that the VT group showed larger improvements during training compared to the ST group and, in fact, the correlation between WM composite score improvement and

improvement during training approached significance ($r = .292$; $p = .054$). These data suggest that perhaps the variability in the literature is due to variability in the efficacy of training on training task performance.

Attentional Control

One of the most consistent findings in the cognitive training literature is that cognitive training transfers to measures of AC. Nearly all studies that assess AC report this positive transfer effect (e.g., Chein & Morrison, 2010; Klingberg, et al., 2002; Olesen, et al., 2004; Westerberg & Klingberg, 2007). In fact, there was only one exception where researchers failed to find training-related improvements on measures of AC (Dahlin, Neely, et al., 2008; Owen, et al., 2010). It is important to note, however, that all of these studies assessed AC with a single task. In fact all studies used the Stroop task. The importance of measuring cognitive processes with multiple measures has been described previously (Morrison & Chein, 2011; Redick, et al., 2011; Shipstead, et al., 2010), and therefore, it is evident that more research is necessary to broaden the scope of training-related improvements on measures of AC.

In the current study, AC was assessed using both the flanker task and the antisaccade task. These tasks were selected because of their high and shared loadings onto an AC construct (e.g., Unsworth & Spillers, 2010). The current antisaccade and flanker data show transfer effects that approach significance. When data were collapsed across the training groups, training effects emerged revealing significant improvement at BS2 for the combined training group compared to the NCC group. In this instance, composite scores were inappropriate as the flanker and antisaccade measures did not correlate. While the correlation between the flanker and antisaccade tasks is often

significant (e.g., $r = -.35$; Unsworth & Spillers, 2010), these data compare flanker task congruency effects to antisaccade task accuracy. This comparison in the current data indicated similar trend that approached significance ($r = -.24$, $p = .062$), however, accuracy performance on the antisaccade task was consistent across groups (all groups improved similarly from BS1 and BS2) and the interesting training effects exist only in the RT data.

Fluid Intelligence

The cognitive training literature is especially inconsistent concerning measures of *Gf*. Approximately half of the studies that assess the effectiveness of cognitive training on measures of *Gf* report significant post-training improvement using both the RAPM and BOMAT tasks (e.g., Colom et al., 2010; Jaeggi, et al., 2008; Jaeggi, et al., 2010; Klingberg, et al., 2002; Olesen, et al., 2004). The other half, however, fail to show an effect of training on *Gf* performance also using the RAPM as well as RAPM, CFF, letter and number sets, inferences, and analogies (e.g., Chein & Morrison, 2010; Dahlin, Nyberg, et al., 2008; Redick, et al., 2011; Schwarb, et al., under review; Westerberg & Klingberg, 2007). The current data lend further support in favor of a lack of effect on *Gf* following WM training. The question then, is obvious, why do some experiments show a positive transfer effect while others do not? Future research is necessary to tease apart the differences both in training tasks used as well as transfer task administration to try to uncover some common features that support transfer. Recently there has been some speculation as to what drives the inconsistent findings, and researchers have concluded that perhaps the amount of training gain is pivotal to successful transfer to measures of *Gf* (Jaeggi, Buschkuhl, Jonides, & Shah, 2011). The current data, however, do not support

this hypothesis as the correlation between *Gf* composite score improvement and improvement during training was small, not significant, and, in fact, negative ($r = -.11$, $p = .470$).

Visual Short-Term Memory

As mentioned in Chapter 4, most studies investigating cognitive training and VSTM use VSTM tasks during training and therefore any transfer to other measures of VSTM only constitutes near-transfer. Olesen and colleagues (2004) were the first to train participants on several WM tasks and investigate the effect of training on an untrained measure of VSTM (i.e., the span board task). These data indicate a trend toward significant WM training transfer to VSTM (i.e., $p = .12$). The current study sought to investigate the role of training in VSTM more rigorously and included three separate measures of VSTM: The change detection task, the color short-term recall task, and the spatial short-term recall task. Furthermore, the current study sought to specifically investigate the effect of training on two subprocesses of VSTM, namely number and resolution.

The current data show no evidence for an effect of WM training on VSTM as measured by either of the short-term recall tasks on either the measure of number (i.e., *Pmem*) or the measure of resolution (i.e., *SD*). The only significant effect was a significant decrease in resolution following training for the VT group compared to both the ST and NCC groups. This effect is surprising and not likely to be the result of WM training, though the alternative is not apparent. VSTM capacity estimates were also extracted from both the short-term recall tasks for each of the three groups. There was no indication that training had any influence on these capacity estimates. Thus it seems that

the training design used in the current study was ineffective in altering VSTM performance as measured by the short-term recall task.

When the change detection data are considered, however, the story changes dramatically. Possible reasons for this discrepancy are discussed in the following section (i.e., What, Then, Does Working Memory Training Train?). These data indicate significant training-related improvements on change detection task performance for both the VT and ST groups consistent with our previous work (Schwarb, et al., under review). Additionally, the current study shows training-related improvement on the number of items held in memory (i.e., accuracy on between-class change trials). Furthermore, individual capacity estimates were extracted revealing a significant increase in capacity following adaptive *n*-back training for the ST group. There was a trend for an improvement in capacity for the VT group as well. Finally, there was no evidence of increased capacity for the NCC group. To the best of my knowledge, this is the first study that has directly assessed the influence of WM training on VSTM capacity estimates.

Finally, the current study also shows significant training-related improvement on measures of resolution (i.e., within-class change trials) in the change detection task for both groups. This finding is somewhat surprising. Studies suggest that VSTM resolution can be altered, but only when participants have expertise with the stimuli (Curby & Gauthier, 2007; Scolari, et al., 2008). In the current study, all three groups had equal exposure to the specific stimuli used in the change detection task; therefore, VT and ST group participants should not have an elevated level expertise compared to NCC group participants. However, improvement is evident and consistent between the training groups. While previous data have suggested that stimulus specific expertise is critical for

resolution enhancements (Curby & Gauthier, 2007; Scolari, et al., 2008), consider a situation where expertise is construed more broadly. Perhaps participants in one or both training groups could acquire expertise with specific perceptual judgments resulting in perceptual learning and improved discrimination among representations in memory. In order to discriminate between items on within-class change trials, participants must be able to identify the component features of the stimuli and their relationship with each other. At the perceptual level, letter discrimination also requires identification of component features and the relationship between them (e.g., Pelli, Burns, Farrell, & Moore-Page, 2006). It is arguable, therefore, that participants in the VT group gained a certain level of expertise with such perceptual judgments spurring improved resolution in the change detection task after training. Such an argument, however, does not translate to participants in the ST group who showed equivalent improvement in resolution after training.

Contact Control Group

In the present study, two experimental groups were designed each to serve as the contact control group for the other. One challenge of selecting an appropriate contact control group is that when different training tasks are used, these tasks often differ in many important ways (e.g., underlying cognitive processes, task demands, and level of difficulty). Furthermore, it is likely that researchers are unaware of all of the levels at which cognitive processing differs between the two tasks. Consequently, interpretation of the data is rarely straight forward as there are typically a number of alternative explanations. Selecting highly similar training tasks, as in the present experiment, is advantageous in that the tasks are matched on both task requirements and difficulty.

However, because the underlying processes are very similar, the likelihood of identifying different training effects between the groups may be reduced. In the present study, a spatial and a verbal version of the adaptive *n*-back task were selected with the hypothesis that there might be stimulus specific benefits of training on some of the battery tasks. For example, if the verbal adaptive *n*-back task trains verbal WM, then one might expect to see a benefit of training on the operation span task and not the symmetry span task; or if the spatial adaptive *n*-back task trains spatial WM, then one might expect to see a benefit of training on the symmetry span task or the spatial short-term recall task, but not the RAPM task. In the present study, such stimulus specific transfer was not evident. Rather, when an effect of training existed, the effect was typically similar for both training groups. There are notable exceptions however, for example, the training benefit on measures of WM (i.e., WM composite scores) was restricted to the VT group. Conversely, there was an ST group advantage on composite measures of VSTM number.

Despite the similarity in ST and VT group results, I do not believe that these data can be consistently interpreted with a Hawthorne effect, demand characteristic, or group motivation argument (cf. Shipstead, et al., 2010). In the present data, there are not universal benefits of training on all tasks measured. Thus it seems unlikely that differences in motivation, expectations, or demand characteristics would have a differential effect on the various battery of tasks used here. Why, for example, might motivational differences or interactions with experimenters positively influence measures of AC, but not measures of *Gf*; or some measures of VSTM (i.e., change detection task), but not other VSTM measures (i.e., short-term recall tasks)? There is no obvious reason to believe that motivation or expectation works selectively among the battery tasks.

What, Then, Does Working Memory Training Train?

The majority of studies in the cognitive training literature that train on the *n*-back task categorize themselves as WM training studies. However, while successful performance on the *n*-back task surely requires WM storage processes (e.g., McElree, 2001; Shipstead, et al., 2010), there are certainly other cognitive processes engaged. For example, AC processes are also essential for accurate performance on the *n*-back task (e.g., Jaeggi, et al., 2008; McElree, 2001; Shipstead, et al., 2010; Verhaeghen, Cerella, & Basak, 2004). The current data suggest that both storage and AC are improved during training with the *n*-back task, and I argue here that it is improved AC processes that are primarily responsible for improvement on the various battery tasks.

Attentional control is the process, or set of processes, that allows an individual to select task/situation relevant stimuli and ignore other stimuli (e.g., Neill, Valdes, & Terry, 1994). In other words, AC facilitates relevant processing while inhibiting irrelevant processing to ensure optimal performance. Attention can select on certain physical attributes of the stimulus such as color and location (e.g., Broadbent, 1958; Neill, et al., 1994). In the spatial version of the adaptive *n*-back task, individual stimuli are differentiated by their spatial location; however, equally critical is each stimulus' position in time. With practice during training, the ability to select on time and space is refined. Important to note, however, is that in the current version of the task fine-grained precision of spatial selection is not paramount as each location is differentiable by several degrees of visual angle. So, it is reasonable to believe that inhibiting stimuli at a distance rather than in close proximity is being trained. In the verbal version of the adaptive *n*-

back task, spatial location is not important; however, again temporal position is critical for successful performance.

Explanation of Failed Transfer

If this hypothesis holds merit, then it is expected that performance on both of the short-term recall tasks as well as the RAPM and CFF tasks would not show a benefit of n -back training. In all of these tasks, all of the stimuli are relevant to responding appropriately. In the short-term recall tasks, all stimuli in the memory set have an equally likely chance of being probed for retrieval, therefore, the best strategy is to select all available stimuli. Thus improving the ability to inhibit irrelevant stimuli is not beneficial to task performance, and consequently performance on both of the short-term recall tasks do not benefit from adaptive n -back training.

The same is true of both measures of Gf tested in this study. In order to correctly identify the missing information in the RAPM task, an individual must pay close attention to all of the other stimuli to try to understand what best completes the pattern. Inhibiting any part of the problem set is not useful because critical information about the best solution is embedded in all of the related stimuli. Furthermore, all of the relevant information for a given problem is available simultaneously, so temporal distinctions are not necessary. In the CCF task, similarly all of the information necessary to solve the problem is presented concurrently and all of the choices must be considered and compared to arrive at the appropriate solution. These tasks, therefore, rely little if at all on the component processes that were improved via single n -back training.

Explanation of Successful Transfer

Each of the complex span tasks used to measure WM required participants to keep track of information which was temporally interleaved with irrelevant information. Thus participants were required to keep track of some information (i.e., the letter in the automated operation span task and the spatial locations in the automated symmetry span task) while inhibiting irrelevant information (i.e., math problems in the automated operation span task and symmetry judgments in the automated symmetry span task) presented just before and after that relevant information. Thus, temporal selection is essential to accurate performance in these tasks. If the hypothesis that improvement on the verbal adaptive *n*-back task across training sessions is due to improved AC in the temporal domain, that it is unsurprising that the VT group alone would show a benefit of training on the composite measure of WM as it is the temporal distinctions that are critical in complex span tasks.

In the antisaccade task, participants must detect a cue positioned to the far right or far left of the screen and then when the cue disappears, participants must focus their attention on the opposite side of the screen in order to detect a target stimulus. Thus, there is both a necessary temporal (i.e., cue occurs prior to the target) and spatial (i.e., shifting attention to the opposite side of space) component to this task. Thus it is perhaps unsurprising that both the VT and ST groups show a benefit of training compared to the NCC groups because both groups can take advantage of improved temporal controlled attention. Furthermore, although the Time x Group interaction was only significant when the data were collapsed across training type, the effect approached significance when the training groups were considered independently. Additionally, these data suggest that the

ST group showed a larger benefit of training than the VT group compared to the NCC group. The hypothesis proposed here also predicts this pattern; it is expected that the ST group would show a larger advantage compared to the VT group because these participants can take advantage of both improved temporal and spatial controlled attention while the VT group participants are only aided by improved temporal AC processes and do not benefit from improved spatial controlled attention as well.

The flanker task also requires the selection of task relevant and the inhibition of task irrelevant information; participants must respond to the direction of a central arrow while ignoring the direction of the flanking arrows. In this task, the target is distinguishable from the distractors based solely on spatial location, thus training that encourages accurate special selection should, in theory, transfer to this task. This perhaps, could explain why participants in the ST group show an advantage of training on the flanker task. This, however, does not explain why VT participants also show a benefit. In the flanker task, however, once the target is identified a correct response requires participants to identify the relationship among similar features shared between the target and distractors (i.e., the letter V rotated to the right (<) versus the letter V rotated to the left (>)). As discussed in the previous section, it seems plausible that during training participants in the VT group are learning to make perceptual judgments requiring the identification of the relationship between component features of the letters. Perhaps then it is this improved ability to identify the relationship among component parts that is aiding performance for the VT group. Therefore, while both groups show improvement after training, performance is enhanced on different task relevant dimensions for each group.

While strong trends for training improvement are evident in the flanker data, given these hypotheses, why are the effects not significant when the training groups are considered separately? I believe that for the ST group, this is likely due to the spatial distance separating targets and distracters in the spatial version of the *n*-back used in the current study. In the adaptive spatial *n*-back task, the display consisted of a 4 x 4 grid that subtended 16° of visual angle. Consequently, from center to center, any two spatial locations were a minimum of 4° of visual angle apart. Conversely, in the flanker task the entire display subtended 5° of visual angle with the relevant and irrelevant stimuli separated only by 1° of visual angle. Accordingly, the flanker task requires a much more spatially local degree of discrimination and this level of precision was not targeted by the training task. The verbal training task, conversely, does not train any spatial component; thus the VT group participants do not benefit from improved spatial discrimination. Once the target is identified, however, performance is enhanced by improved feature discrimination facilitating a quick response. Thus, it seems like both the verbal and spatial training tasks could be better tailored to facilitate successful transfer to the flanker task. Future research is necessary to explore this issue.

Finally, in the change detection task, participants are cued to one or the other side of the computer screen and are responsible for remembering the stimuli that appear on that side of space. Inhibiting the stimuli on the uncued side of space is essential to accurate performance. Indeed past research has demonstrated the importance of inhibiting irrelevant stimuli in the change detection task. Vogel, McCollough, and Machizawa (2005) asked participants to complete a change detection task where they were presented with red and blue bars at various orientations and were instructed to pay attention only to

the red bars in the cued hemifield while ERPs were recorded from the scalp. Participants were divided into high and low WM capacity groups based on change detection task performance. Interestingly, this CDA amplitude (focused over the posterior parietal and lateral occipital electrode sites) was found to be highly sensitive to the number of items in the memory array and reached asymptote at around four items. CDA amplitude was compared for each group in three conditions: 2 red bars, 4 red bars, and 2 red and 2 blue bars (distractor condition). For the high capacity group, when participants only had to remember the orientation of two bars (i.e., 2 red bars condition vs. distractor condition), CDA amplitudes were similar and less than when participants had to remember the orientation of four bars (i.e., 4 red bar condition). For the low capacity group, however, CDA amplitude was similarly high for the 4 red bars condition and distractors condition indicating that participants were unable to ignore the irrelevant information in the distractor display. Thus, it seems that high capacity individuals may be able to better focus their attention on the task relevant stimuli. Important to note, however, is a major difference between the current study and these data is that participants were required to inhibit information both from the uncued as well as the cued side of the display.

So in the change detection task, after an individual has inhibited the irrelevant side of space and encoded the proper stimuli, a short pause occurs and a probe stimulus is presented. Participants must determine whether or not the probe is identical to the stimulus presented one second previously and now held in memory. Thus, the change detection task requires both spatial and temporal AC to accurately select the appropriate representations to perform accurately. Therefore, it is perhaps unsurprising that both participants in the VT and ST groups show a post-training benefit. It is perhaps surprising

that VT and ST group participants show equivalent improvement with training, however, the design of the current study does not allow us to identify the potentially unique contributions of spatial and temporal AC for each group on this task and future research is necessary to tease these data apart. Furthermore, both training groups show similar post-training improvement on measures of number and resolution. If we believe these processes to be independent from each other (Awh, et al., 2007), then perhaps there is little reason to believe that training would influence performance on both processes in a similar way as indicated by the current data. Thus again further research is necessary to understand the details of these findings.

Cross-Trial Attentional Control

This explanation of results thus far has focused on within-trial inhibition of irrelevant information. This explanation dichotomizes the tasks into those that include irrelevant information on each individual trial and those that do not. While inhibition certainly exists at the within-trial inter-item level, cross-trial inhibition is likely also important. Indeed the *n*-back task is a continuous processing task and therefore does not have independent trials. On each trial, participants must remember new information and inhibit previously learned information that is no longer task relevant. If this type of cross-trial inhibition is improved by *n*-back training, one might expect to see evidence of improved cross-trial inhibition in a task with discrete trials as well.

It is likely that cross-trial interference existed in most, if not all, of the battery tasks used in the current study. With the exception of the *Gf* tasks, stimuli were consistent across trials and therefore previous trial stimuli likely interfered with current trial to some degree. For the *Gf* tasks, participants may have applied similar strategies across trials

which likely caused interference when a new strategy was required. Therefore, if cross-trial AC improves with *n*-back training, one might expect to see post-task improvement on all of the battery tasks.

Cross-trial AC can be assessed in the current study. A common finding using the flanker task is the existence of sequential effects; or rather smaller differences between incongruent and congruent trials following incongruent trials compared to following congruent trials (e.g., Gratton, Coles, & Donchin, 1992; Stümer, Leuthold, Soetens, Schröter, & Sommer, 2002). This characteristic effect has been explained as a tightening of AC following incongruent trials and a loosening of attention after congruent trials (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Therefore, if AC improves after *n*-back training, then perhaps participants would be able to focus (or tighten) their attention on the task relevant stimulus (i.e., the central arrow) and the sequential effect would be reduced as participants would experience less interference from the previous trial.

To test this hypothesis, sequential effects on the flanker task were extracted for each participant (Figure 16). Data were then subjected to a Group (VT, ST, NCC) x Previous Trial Congruency (congruent, incongruent) x Current Trial Congruency (congruent, incongruent) repeated measures ANOVA. The critical Group x Previous Trial Congruency x Current Trial Congruency was significant, $F(2,61) = 4.40$, $p = .016$, $\eta_p^2 = .13$. Session (BS1, BS2) x Previous Trial Congruency (congruent, incongruent) x Current Trial Congruency (congruent, incongruent) repeated measures ANOVAs were conducted for each group individually to understand this significant interaction. The VT and ST groups failed to show a significant Session x Previous Trial Congruency x Current Trial Congruency interaction, $F(1, 19) = 1.75$, $p = .212$, $\eta_p^2 = .08$ and $F(1,21) = .01$, $p = .931$,

$\eta_p^2 = 0$ respectively, indicating that the sequential effect were similar before and after training. The NCC group, however, did show a significant Session x Previous Trial Congruency x Current Trial Congruency interaction, $F(1, 21) = 5.17, p = .034, \eta_p^2 = .20$, demonstrating a significantly smaller sequential effect during BS2 compared to BS1. Taken together, these data indicate a unique reduction in sequential effects in the absence of training (i.e., a reduced sequential effect for the NCC group only). These data are unexpected and also inconsistent with the notion that training-related improved cross-trial AC is influencing post-training performance on untrained cognitive tasks. Further investigation of this issue is certainly necessary to draw definitive conclusions.

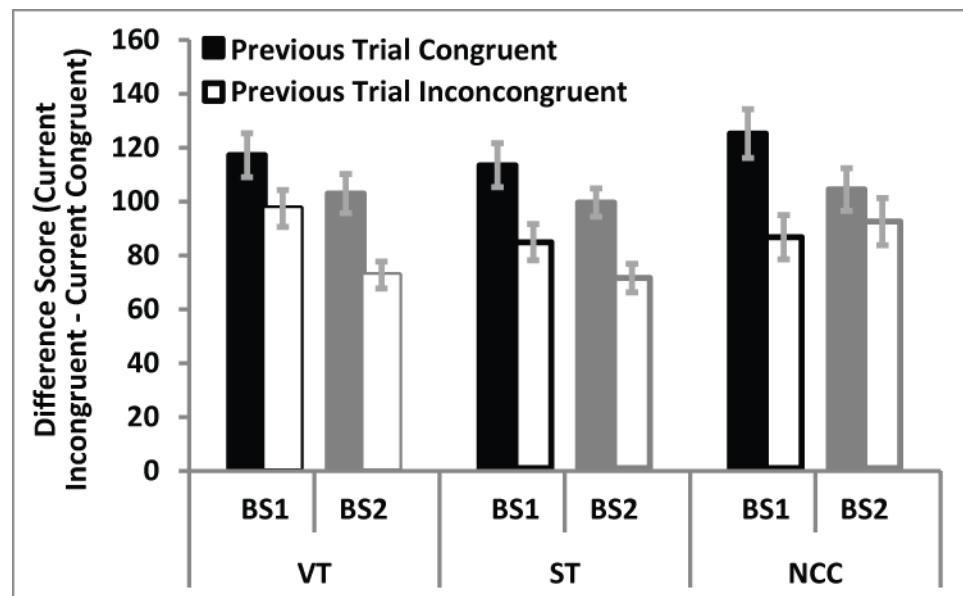


Figure 16. Flanker Task Sequential Effects

Conclusions

Three main conclusions can be drawn from these data. First, an optimized training design does, in fact, promote a high level of performance across training trials. This may be important if the amount of improvement demonstrated during training is related to the amount of improvement on other untrained tasks (i.e., the WM data presented here). Second, adaptive *n*-back training can be effective in improving performance on related, but separate tasks; however, transfer is not global and only occurs when the processes that improved during training are also required in the transfer tasks (e.g., Dahlin, Neely, et al., 2008; Dahlin, Nyberg, et al., 2008). Finally, WM training is an effective means of improving performance on the change detection task and training influences both number and resolution processes. To the best of my knowledge, this is the first study demonstrating significant improvement on a measure of VSTM number and resolution following WM training. Recently inconsistencies in the literature and flaws in experimental design have lead researchers to question the importance or viability of brain training to enhance performance on untrained tasks (Owen, et al., 2010; Shipstead, et al., 2010). The current study suggests that perhaps this situation might not be so dire and that with thoughtful experimental design and reasonable expectations about the breadth of transfer, perhaps a coherent and informative literature can develop providing a comprehensive and comprehensible foundation for cognitive improvement.

REFERENCES

- Adams, J. A. (1987). Historical review and appraisal of research on the learning, retention, and transfer of human motor skills. *Psychological Bulletin*, 101(1), 41-74.
- Alvarez, G. A., & Cavanagh, P. (2004). The capacity of visual short-term memory is set both by visual information load and by number of objects. *Psychological Science*, 15(2), 106-111.
- Awh, E., Barton, B., & Vogel, E. K. (2007). Visual working memory represents a fixed number of items regardless of complexity. *Psychological Science*, 18(7), 622-628.
- Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple-component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28-61). Cambridge: Cambridge University Press.
- Bahrick, H. P. (1984). Semantic memory content in permastore: Fifty years of memory for spanish learned in school. *Journal of Experimental Psychology: General*, 113(1).
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin*, 128(4), 612-637.
- Bartlett, F. C. (1947). The measurement of human skill. *British Medical Journal*, 1(4511), 877-880.
- Beck, M. R., Angelone, B. L., Levin, D. T., Peterson, M. S., & Varakin, A. (2008). Implicit learning for probable changes in a visual change detection task. *Consciousness and cognition*, 17, 1192-1208.
- Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, 129(3), 387-398.
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, 108, 624-652.
- Broadbent, D. E. (1958). *Perception and communication*. New York: Pergamon Press.

- Burton, H., Snyder, A. Z., Conturo, T. E., Akbudak, E., Ollinger, J. M., & Raichle, M. E. (2002). Adaptive changes in early and late blind: A fMRI study of Braille reading. *Journal of Neurophysiology*, 87, 589-607.
- Cattell, R. B. (1949). *Culture free intelligence test, scale 1, Handbook*. Champaign, Illinois: Institute of Personality and Ability.
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects within a complex working memory span task. *Psychonomic Bulletin & Review*, 17(2), 193-199.
- Cogmed. (2006). Cogmed Working Memory Training: Cogmed America Inc.
- Colom, R., Quiroga, M. A., Shih, P. C., Martinez, K., Burgaleta, M., Martinez-Molina, A., . . . Ramirez, I. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence*, 38, 497-505.
- Conway, A. R. A., Cowan, N., & Bunting, M. F. (2001). The cocktail party phenomenon revisited: The importance of working memory capacity. *Psychonomic Bulletin & Review*, 8(2), 331-335.
- Cowan, N. (1999). An embedded-processes model of working memory. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 62-101). Cambridge: Cambridge University Press.
- Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.
- Curby, K. M., & Gauthier, I. (2007). A visual short-term memory advantage for faces. *Psychonomic Bulletin & Review*, 12, 1127-1133.
- Dahlin, E., Neely, A. S., Larsson, A., Backman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science*, 320, 1510-1512.
- Dahlin, E., Nyberg, L., Backman, L., & Neely, A. S. (2008). Plasticity of executive functioning in young and older adults: Immediate training gains, transfer, and long-term maintenance. *Psychology and Aging*, 23(4), 720-730.
- Dempster, F. N. (1988). A case study in the failure to apply the results of psychological research. *American Psychologist*, 43(8), 627-634.
- Ebbinghaus, H. E. (1913). *Memory: A contribution to experimental psychology* (Ruger, H. A. & Bussenius, C. E., Trans.). New York: Dover. (Original work published 1885)

- Eng, H. Y., Chen, D., & Jiang, Y. H. (2005). Visual working memory for simple and complex visual stimuli. *Psychonomic Bulletin & Review*, 12(6), 1127-1133.
- Engle, R. W. (2001). What is working-memory capacity? In H. L. Roediger & J. S. Nairne (Eds.), *The Nature of Remembering: Essays in Honor of Robert G. Crowder* (pp. 297-314). Washington, D.C.: American Psychological Association.
- Engle, R. W., Kane, M. J., & Tuholski, S. W. (1999). Individual differences in working memory capacity and what they tell us about controlled attention, general fluid intelligence, and functions of the prefrontal cortex. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 102-134). Cambridge: Cambridge University Press.
- Erickson, K. I., Colcombe, S. J., Wadhwa, R., Bherer, L., Peterson, M. S., Scalf, P. E., . . . Kramer, A. F. (2006). Training-induced functional activation changes in dual-task processing: An fMRI study. *Cerebral Cortex*, 17, 192-204.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143-149.
- Ester, E. F., Serences, J. T., & Awh, E. (2009). Spatially Global Representations in Human Primary Visual Cortex during Working Memory Maintenance. *Journal of Neuroscience*, 29(48), 15258-15265.
- Fougnie, D., Asplund, C. L., & Marois, R. (2010). What are the units of storage in visual working memory? *Journal of Vision*, 10(12), 1-11.
- Fukuda, K., Awh, E., & Vogel, E. K. (2010). Discrete capacity limits in visual working memory. [Article]. *Current Opinion in Neurobiology*, 20(2), 177-182.
- Fukuda, K., & Vogel, E. K. (2009). Human variation in overriding attentional capture. *Journal of Neuroscience*, 29(27), 8726-8733.
- Fukuda, K., & Vogel, E. K. (2011). Individual differences in recovery time from attentional capture. *Psychological Science*, 22(3), 361-368.
- Fukuda, K., Vogel, E. K., Mayr, U., & Awh, E. (2010). Quantity not quality: The relationship between fluid intelligence and working memory capacity. *Psychonomic Bulletin & Review*, 17(5), 673-679.
- Gao, Z., Li, J., Liang, J., Chen, H., Yin, J., & Shen, M. (2009). Storing fine detailed information in visual working memory—Evidence from event-related potentials. *Journal of Vision*, 9(7), 1-12.

- Garavan, H., Kelley, D., Rosen, A., Rao, S. M., & Stein, E. A. (2000). Practice-related functional activation changes in a working memory task. *Microscopy Research and Technique*, 51, 54-63.
- Gratton, G., Coles, M. G. H., & Donchin, O. (1992). Optimizing the use of information: Strategic control of activation and responses. *Journal of Experimental Psychology: General*, 121, 480-506.
- Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, 423, 534-537.
- Green, C. S., & Bavelier, D. (2006). Effect of action video games on the spatial distribution of visual attention. *Journal of Experimental Psychology: Human Learning and Memory*, 32(6), 1465-1478.
- Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance*, 28(3), 527-545.
- Heitz, R. P., & Engle, R. W. (2007). Focusing the spotlight: Individual differences in visual attention control. *Journal of Experimental Psychology: General*, 136(2), 217-240.
- Heuer, H., & Schmidtke, V. (1996). Secondary-task effects on sequence learning. *Psychological Research/Psychologische Forschung*, 59(2), 119-133.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, 105(19), 6829-6833.
- Jaeggi, S. M., Buschkuhl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, 108(25), 11081-11086.
- Jaeggi, S. M., Studer-Luethi, B., Buschkuhl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning -- implications for training and transfer. *Intelligence*, 38(6), 625-635.
- James, W. (1890). *The principles of psychology* (Vol. 1). New York: Holt.
- Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, 59, 193-224.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: An experience-sampling study of

- working memory and executive control in daily life. *Psychological Science*, 18, 614-621.
- Kane, M. J., Conway, A. R. A., Miura, T. K., & Colflesh, G. J. H. (2007). Working memory, attentional control, and the n-back task: A question of construct validity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(3), 615-622.
- Kane, M. J., & Engle, R. W. (2003a). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47-70.
- Kane, M. J., & Engle, R. W. (2003b). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to Stroop interference. *Journal of Experimental Psychology: General*, 132(1), 47-70.
- Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, 377, 155-158.
- Kerr, R., & Booth, B. (1978). Specific and varied practice of motor skill. *Perceptual and motor skills*, 46(2), 395-401.
- Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Increased brain activity in frontal and parietal cortex underlies the development of visuospatial working memory capacity during childhood. *Journal of Cognitive Neuroscience*, 14(1), 1-10.
- Li, S. C., Schmiedek, F., Huxhold, O., Rocke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging*, 23(4), 731-742.
- Luck, S. J., & Vogel, E. K. (1997). The capacity of visual working memory for features and conjunctions. [Article]. *Nature*, 390(6657), 279-281.
- Luria, R., Sessa, P., Gotler, A., Jolicoeur, P., & Dell'Acqua, R. (2009). Visual short-term memory capacity for simple and complex objects. *Journal of Cognitive Neuroscience*, 22(3), 496-512.
- Maguire, E. A., Gadian, D. G., Johnsrude, I. S., Good, C. D., Ashburner, J., Frackowiak, R. S. J., & Frith, C. D. (2000). Navigation-related structural changes in the hippocampi of taxi drivers. *Proceedings of the National Academy of Sciences*, 97(8), 4398-4403.
- McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 817-835.

- Mercado, E. (2008). Neural and cognitive plasticity: From maps to minds. *Psychological Bulletin*, 134(1), 109-137.
- Miller, G. A. (1956). The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63, 81-96.
- Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence*, 37, 327-328.
- Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychological Bulletin & Review*, 18, 46-60.
- Neill, W., Valdes, L., & Terry, K. (1994). Selective attention and the inhibitory control of cognition. In F. N. Dempster & C. Brainerd (Eds.), *New perspectives on interference and inhibition in cognition*. New York: Academic Press.
- Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, 7(1), 75-79.
- Olson, I. R., & Jiang, Y. H. (2004). Visual short-term memory is not improved by training. *Memory & Cognition*, 32(8), 1326-1332.
- Olson, I. R., Jiang, Y. H., & Moore, K. S. (2005). Associative learning improves visual working memory performance. *Journal of Experimental Psychology: Human Perception and Performance*, 31(5), 889-900.
- Owen, A. M., Hampshire, A., Grahn, J. A., Stenton, R., Dajani, S., Burns, A. S., . . . Ballard, C. G. (2010). Putting brain training to the test. *Nature*, 465(7299), 775-779.
- Pashler, H. (1988). Familiarity and visual change detection. *Perception & Psychophysics*, 44(4), 369-378.
- Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review*, 14(2), 187-193.
- Pelli, D. G., Burns, C. W., Farrell, B., & Moore-Page, D. C. (2006). Feature detection and letter identification. *Vision Research*, 46, 4646-4674.
- Petersen, S. E., van Mier, H., Fiez, J. A., & Raichle, M. E. (1998). The effects of practice on the functional anatomy of task performance. *Proceedings of the National Academy of Sciences*, 95, 853-860.

- Posner, M. I., DiGirolamo, G. J., & Fernandez-Duque, D. (1997). Brain mechanisms of cognitive skill. *Consciousness and cognition*, 6, 267-290.
- Raven, J. C. (1990). *Advanced Progressive Matrices. Sets I, II*. Oxford: Oxford University Press.
- Redick, T. A., Hambrick, D. Z., Shipstead, Z., Harrison, T. L., Hicks, K. L., & Engle, R. W. (2011). *Generalization of working memory training to complex cognition?* Paper presented at the Psychonomic Society Annual Meeting, Seattle, WA.
- Rogers, W. A. (1996). Assessing age-related differences in the long-term retention of skills. In W. A. Rogers, A. D. Fisk & N. Walker (Eds.), *Aging and skilled performance: Advances in theory and application* (pp. 185-200). Hillsdale, NJ: Erlbaum.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Paradigms suggest new concepts for training. *Psychological Science*, 3(4), 207-217.
- Schmiedek, F., Lövdén, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO study *Frontiers in Aging Neuroscience*, 2(27), 1-10.
- Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors*, 27(3), 285-300.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.
- Schumacher, E. H., Hendricks, M. J., & D'Esposito, M. D. (2005). Sustained involvement of a frontal-parietal network for spatial response selection with practice of a spatial choice-reaction task. *Neuropsychologia*, 43(10), 1444-1455.
- Schwarb, H., Nail, J., McKinley, R. A., Tripp, L., Keilholz, S., & Schumacher, E. H. (under review). Moderate amounts of spatial and verbal n-back training improves working memory capacity as measured by untrained working memory tasks.
- Scolari, M., Vogel, E. K., & Awh, E. (2008). Perceptual expertise enhances the resolution but not the number of representations in working memory. *Psychonomic Bulletin & Review*, 15(1), 215-222.
- Serences, J. T., Ester, E. F., Vogel, E. K., & Awh, E. (2009). Stimulus-specific delay activity in human primary visual cortex. *Psychological Science*, 20(2), 207-214.
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, 5(2), 179-187.

- Shipstead, Z., Redick, T. A., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica*, 50, 245-276.
- Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs*, 74(11), 1-29.
- Stürmer, B., Leuthold, H., Soetens, E., Schröter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, 28, 1345-1363.
- Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions: II. The estimation of magnitudes. *Psychological Review*, 8, 553-564.
- Todd, J. J., & Marois, R. (2004). Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature*, 428(6984), 751-754.
- Todd, J. J., & Marois, R. (2005). Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. [Review]. *Cognitive Affective & Behavioral Neuroscience*, 5(2), 144-155.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory & Language*, 28, 127-154.
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. [Research Support, U.S. Gov't, Non-P.H.S.]. *Behavior Research Methods*, 37(3), 498-505.
- Unsworth, N., Schrock, J. C., & Engle, R. W. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1302-1321.
- Unsworth, N., & Spillers, G. J. (2010). Working memory capacity: Attention control, secondary memory, or both? A direct test of the dual-component model. *Journal of Memory & Language*, 62, 392-406.
- Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to expand the focus of serial attention from one to four items in 10 hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10(6), 1322-1337.
- Vogel, E. K., & Machizawa, M. G. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature*, 428(6984), 748-751.

- Vogel, E. K., McCollough, A. W., & Machizawa, M. G. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature*, 438(7067), 500-503.
- Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory — a single-subject analysis. *Physiology & Behavior*, 92, 186-192.
- Willis, S. L., Tennstedt, S. L., Marsiske, M., Ball, K., Elias, J., Koepke, K. M., . . . Wright, E. (2006). Long-term effects of cognitive training on everyday functional outcomes in older adults. *Journal of the American Medical Association*, 296(23), 2805-2814.
- Xu, Y. D., & Chun, M. M. (2006). Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*, 440(7080), 91-95.
- Zhang, W., & Luck, S. J. (2008). Discrete fixed-resolution representations in visual working memory. *Nature*, 453(7192), 233-235.
- Zhang, W., & Luck, S. J. (2009). Sudden death and gradual decay in visual working memory. *Psychological Science*, 20(423-428).
- Zhang, W., & Luck, S. J. (2011). The number and quality of representations in working memory. *Psychological Science*, 22(11), 1434-1441.